

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OW protein - protein search, using sw model

Run on: May 8, 2002, 12:13:32 ; Search time 28.7 Seconds

(Without alignments)
2293.469 Million cell updates/sec

Title: US-09-647-522-5

Sequence: 1 MILKHPMLFIVLATSAXH.....SGYRSMENHPGQGHWSIT 450

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_17:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.potent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2322	100.0	450	5	09GV72 carypdae ra
2	992.5	42.7	463	5	09GNN8 carypdae al
3	137	5.9	1222	2	09X349 bacillus an
4	117	5.0	833	11	062187 mus musculu
5	116	5.0	530	2	09X0M7 thermotoga
6	113.5	4.9	306	2	09R9R9 aeromonas p
7	112.5	4.8	510	10	09S3X2 arabisidopsis
8	112.5	4.8	6713	2	09G9F54 staphylococ
9	112	4.8	788	2	09F5M7 rhizobium m
10	112	4.8	2026	2	09J3P4 neisseria m
11	111.5	4.8	656	2	09W4X0 thermotoga
12	111.5	4.8	1336	5	09W4X0 dictyostell
13	111	4.8	865	2	09XDU1 clostridium
14	111	4.8	1657	5	09B5J9 leptomomas
15	110.5	4.8	656	2	09W4X0 thermotoga
16	110.5	4.8	661	2	09X020 thermotoga
17	110	4.7	810	1	P82857 halobacteri
18	109.5	4.7	874	2	09G277 haemophilus
19	109	4.7	541	2	09HUP8 pseudomonas

20	109	4.7	612	2	09RXY1	09RX1 delinococcus
21	108.5	4.7	659	2	09KVP3	09KVP3 vibrio chol
22	107.5	4.6	387	2	086266	086266 pseudomonas
23	106.5	4.6	1098	2	050733	050733 borrelia bu
24	106	4.6	1369	5	09NB29	09NB29 caenorhabdi
25	106	4.6	1975	2	09K087	09K087 neisseria m
26	105.5	4.5	387	2	086265	086265 pseudomonas
27	105.5	4.5	2504	2	085160	085160 photorhabdu
28	105	4.5	384	2	005492	005492 bacillus ce
29	105	4.5	1456	12	09IMP2	09IMP2 potato viru
30	104.5	4.5	333	2	P72126	P72126 pseudomonas
31	104.5	4.5	406	2	09R67	09R67 bacillus ce
32	104.5	4.5	532	2	09A46	09A46 caudobacter
33	104.5	4.5	739	2	044385	044385 agrobacteri
34	104.5	4.5	1241	12	085444	085444 diadromus p
35	103.5	4.5	349	2	09S000	09S000 escherichia
36	103.5	4.5	440	2	09FC04	09FC04 pseudomonas
37	103.5	4.5	565	2	09S073	09S073 escherichia
38	103.5	4.5	706	9	038305	038305 lactococcus
39	103.5	4.5	859	11	09UKK5	09UKK5 mus musculu
40	103	4.4	504	2	09A8F5	09A8F5 caudobacter
41	103	4.4	638	2	056616	056616 vibrio chol
42	103	4.4	949	11	09DBP9	09DBP9 mus musculu
43	102.5	4.4	460	5	017836	017836 caenorhabdi
44	102.5	4.4	734	5	09BIU2	09BIU2 euagrus chl
45	102	4.4	413	2	085779	085779 myxococcus

ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	450 AA.
ID	09GV72	09GV72		
AC	09GV72	09GV72		
DT	01-MAR-2001	(TREMBLrel. 16, Created)		
DT	01-MAR-2001	(TREMBLrel. 16, Last sequence update)		
DT	01-MAR-2001	(TREMBLrel. 16, Last annotation update)		
DE	TOXIN-1.			
GN	CRT-1.			
OS	Carypdae rastoni.			
OC	Eukaryota; Metazoa; Cnidaria; Cubozoa; Cubomedusae; Carypidae;			
OX	NCBI_TaxID=78582;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20422301; PubMed=10964707;			
RA	Nagai H., Takuwa K., Nakao M., Ito E., Miyake M., Noda M.,			
RA	Nakajima T.;			
RT	"Novel protease from the box jellyfish (sea wasp) Carypdae			
RT	rastoni.			
RT	Biotech. Biochem. Commun. 275:582-588(2000);			
DR	EMBL; AB015878; BAB12728.1;			
SO	SEQUENCE 450 AA; 49392 MW; CD393CF25BFD2FD CRC64;			
Query Match	100.0%; Score 2322; DB 5; Length 450;			
Best Local Similarity	100.0%; Pred. No. 1e-161;			
Matches 450; Conservative	0; Mismatches 0; Indels 0; Gaps 0;			
Db	1 MILKHPMLFIVLATSAXHGRSDVNSLTTRVETALKESGSEALALEGKGEIOT 60			
Qy	1 KPDVGOATITLSSVGSALCLKNSGDATITISGCLDIVAGIATFFGPGVGICAVASV 120			
Db	61 KPDVGOATITLSSVGSALCLKNSGDATITISGCLDIVAGIATFFGPGVGICAVASV 120			
Qy	121 SSILSFTGSSAKNSVAVIDRALSKRDEAIOHHAAGAKRDEFASSAFIOVMQOQSLT 180			
Db	121 SSILSFTGSSAKNSVAVIDRALSKRDEAIOHHAAGAKRDEFASSAFIOVMQOQSLT 180			
Qy	181 DSPLSITAAVNPVYKFSNFIGOLESRISGGAATITSLSDAKRAVDITLLYCOLVYMETLL 240			

```

Db 181 DSDLSIIAANVPYKFSNFGLESRLSOGAATSLSDAKRAVDLILYCOLVYMRTEL 240
QY 241 VDLATILYKGNNAHVASAVNANRVKLEADLDFLHKILIPQALIGAYHHSSETS 300
Db 241 VDLATILYKGNNAHVASAVNANRVKLEADLDFLHKILIPQALIGAYHHSSETS 300
QY 301 KALINTYKFGVDPVPRPIGNRRYKFTNSYWNYSICSEAYMGNYMFRGCSNVRNPIRY 360
Db 301 KALINTYKFGVDPVPRPIGNRRYKFTNSYWNYSICSEAYMGNYMFRGCSNVRNPIRY 360
QY 361 SKMSDGYTEMNSDRKRLYTKHDQGWGCTLDEDPDQGMHFIPLRHGKYVSSKRW 420
Db 361 SKMSDGYTEMNSDRKRLYTKHDQGWGCTLDEDPDQGMHFIPLRHGKYVSSKRW 420
QY 421 NMFMYESSASGYIRSMENNPQGHWSIT 450
Db 421 NMFMYESSASGYIRSMENNPQGHWSIT 450

```

```

RESULT 2
O9GNB8 PRELIMINARY: PRT: 463 AA.
AC 09GNB8;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE TOXIN-A PRECURSOR.
GN CAT-1.
OS Carybdea alata.
OC Eukaryota; Metazoa; Cnidaria; Cubozoa; Cubomedusae; Carybdeidae;
OC Carybdea.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-20422302; Pubmed-10964708;
RT Nagai H., Takawa K., Nakao M., Sakamoto B., Crow G.L., Nakajima T.;
RT Isolation and characterization of a novel protein toxin from the
RT Hawaiian box jellyfish (sea wasp) Carybdea alata.
RL Biochem. Biophys. Res. Commun. 275:589-594(2000).
DR EMBL: AB036714; BAB12727.1; -.
KW Signal.
FT SIGNAL 1 18 POTENTIAL.
SQ SEQUENCE 463 AA; 51604 MW; 461253DF6CE9D3C0 CRC64;

```

```

Query Match 42.7%; Score 992.5; DB 5; Length 463;
Best Local Similarity 43.5%; Pred. No. 1.8e-64;
Matches 197; Conservative 95; Mismatches 148; Indels 13; Gaps 7;

```

```

QY 9 LFLVLAITSAGK-----KSDVNSLITKVEITALKEAS--GSNEAL-ELALEGLKGEIOT 60
Db 11 LFLVLAITSAGKRLSRYSRSDADAVTDIDGITGOLINDIGTITKRLKELAGVQEVK 70
QY 61 KP-DRYGOATKILGSVSGALGKLSNGDPAKIIISGLDIYAGIATTEGGPYMGIGAVASF 119
Db 71 EPATITAKVSTIVSGVSGSLKFKSGDPFVAGSCDIIIVATTEGGPGIAGVAVSL 130
QY 120 VSSITLSLFGSSAKNSAAYIDRALSKRDEALORHAAKRPFAESSAFIOVMKQOQSL 179
Db 131 ISSITLSLFGSSAKNSAAYIDRALSKRDEALORHAAKRPFAESSAFIOVMKQOQSL 190
QY 180 TDSDLSTIANVVPYKFSNFGLESRLSOGAATSLSDAKRAVDLILYCOLVYMRTEL 239
Db 191 TEHLDSVRRAVDATFNMLGYLESLKINGSVTDNNEAMRTINFLTLQLSVARETL 250
QY 240 LVDLATILYK--GNAEHVASAVNANRVKLEADLDFLHKILIPQALIGAYHHSSETS 297
Db 251 LFGVILYKAGAGAYDELALSLTSDQNKRETRVTFLOHMETKYSICGSSYYPIDHS 310
QY 298 ETSKALINTYKFGVDPVPR-PIGNRRYKFTNSYWNYSICSEAYMGNYMFRGCSNVRNP 356
Db 311 KALIGILKILKFGVDPVPRPIGNRRYKFTNSYWNYSICSEAYMGNYMFRGCSNVRNP 370

```

```

QY 357 NFRVSKMSDGYTEMNSDRKRLYTKHDQGWGCTLDEDPDQGMHFIPLRHGKYVSS 416
Db 371 GIRIKLENGYHTI-ILRSKAMVYTHAOGMGWGTDEDEPGEQYTFPIPLTNGFIWST 429
QY 417 KRPWPMFYMESSASGYIRSMENNPQGHWSI 449
Db 430 KRPWPMFYMESSASGYIRSMENNPQGHWSI 462

```

```

RESULT 3
O9X349 PRELIMINARY: PRT: 1222 AA.
AC 09X349;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DE PXO1-79.
OS Bacillus anthracis.
OC Plasmid virulence plasmid PXO1.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-STERNE;
RA Okinaka R.T., Cloud K., Hamton O., Hoffmaster A., Hill K.K., Kelm P.,
RA Koehler T., Lamke G., Kumano S., Mahillon J., Manter D., Martinez Y.,
RA Rieke D.O., Svensson R., Jackson P.J.;
RT "The sequence and organization of PXO1, the large Bacillus anthracis
RT plasmid harboring the Anthrax toxin genes."
RL J. Bacteriol. 0:0-0(1999).
DR EMBL: AF065404; AAD32383.1; -.
KW Plasmid.
SQ SEQUENCE 1222 AA; 130407 MW; A7C117874D097E5B CRC64;

```

```

Query Match 5.9%; Score 137; DB 2; Length 1222;
Best Local Similarity 23.0%; Pred. No. 0.3;
Matches 107; Conservative 55; Mismatches 173; Indels 130; Gaps 23;

```

```

QY 21 GKRSDV-NSLITKVEITALKEASGNEALEGLGEGIOTKPRVGOATKIL--GSVGS 77
Db 310 GVRDDYRNKSKVPRDMDKDEK--TKEDADNPXALCANIGTIDGTSTRAARMKAGQVGS 367
QY 78 ALGKLSGDAKTIISGLDIYAGIATTEGGPYMGIGAVASFVSSITLSLFGSSAKNSVA 137
Db 368 ALGKATTTG-----LAGMA--GAGLSPGAMVAGSOGSALGAAPGAAGRSVA 413
QY 138 AVIDR--ALSKHRDEALORHAAKRPFAESSAFIOVMKQOQSLTDSLS--ITAAVNPV 193
Db 414 AVEGAVNALGKHIGSSIKKAGDYISNLNGES-----YPLTDEBIDNODLATDFEN 464
QY 194 YKFSNFGLESRLSOGAATSLSD-----AKRAVDLILYCOLVYMRTEL 243
Db 465 WKADNPNNSAVASLAKOAFNPASDAEIAKVAKNSQMSFTQROKODLQNMKT----- 519
QY 244 ALYKRGNAEHVASAVNANRVKLEADLDFLHKILIPQALIGAYHHSSETS-- 301
Db 520 ATPY--GNARDLVNAATNA--FQKGYEGDKDTFMSQLPEN-----MSAEKEKQW 566
QY 302 -----ALLNY-----TKYFGV-DVPRPIGNRRYKFTNSYWNYSICSEAYMGNYM 346
Db 567 NDLNLNKGQFRNHAQATTKAGAMVDAKDKQGN--LFDKSYVKKDAFAQLALG--- 621
QY 347 FRCGSNVRNPNITVSKMSGF--TYMENSDR-----RKLYITKHDQGW 368
Db 622 -----KVGRAVEGVKGETLESCHONNGVQALVGASTAFAKKGYYTADHRAGF- 668
QY 389 WGTLDDEPDQGMHFIPLRHGKYVSSKRWPMFYMESSASGY 433
Db 669 ---MKOPAD-----MSQREBAAMKHLDOVKYOGF 696

```

RESULT 4
ID 062187 PRELIMINARY; PRT; 833 AA.
AC 062187;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE TRANSCRIPTION TERMINATION FACTOR 1 (TRANSCRIPTION FACTOR).
GN TTF1 OR TTF-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_Taxid=10990;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ERLICH ASCITES;
MEDLINE=95237206; PubMed=7720715;
RA Evers R., Smid A., Rudloff U., Lottspeich F., Guntt I.;
RT "Different domains of the murine RNA polymerase I-specific termination factor MTF-1 serve distinct functions in transcription termination.";
RL EMBL J. 14:1248-1256(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ERLICH ASCITES;
MEDLINE=95320168; PubMed=7597036;
RA Evers R., Guntt I.;
RT "Molecular coevolution of mammalian ribosomal gene terminator sequences and the transcription termination factor TTF-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:5827-5831(1995).
DR EMBL: X83974; CAA58808.1; -
DR MGD: MGI:105044; Ttfl.
DR InterPro: IPR001005; Myb_DNA_bnd.
DR Pfam: PF00249; myb_DNA-binding; 2.
DR SMART: SM00395; SANT; 2.
DR PROSITE: PS00090; MYB_3; 1.
KW Transcription termination.
SQ SEQUENCE 833 AA; 94534 MW; DEB9C8DC6D3EC6CE CRC64;

Query Match 5.0%; Score 117; DB 11; Length 833;
Best Local Similarity 22.0%; Pred. No. 4.8;
Matches 100; Conservative 54; Mismatches 127; Indels 174; Gaps 23;

QY 4 KHL-----PMLFYLAITSKHKRSDVNSLLTYETALKEASG-SNEALEALEGLK--- 55
DB 235 KHLQKAPW-----DVGQSQPEPISLPPEPLSSDELEKSTEAFAVFCRSLKKNV 286
QY 56 ---GEIOTKPDYRGATKILGVSALGKLNCGDATKIISGCLDIYAGIATTFGSPVGW 112
DB 287 FNSQLEPIPD-----SLDDETSISERLD-----STHGGANVAG 321
QY 113 -----IGAVASFVSSILSLFTSSAKNSVAVIDR----- 142
DB 322 ECESTKSHSIKKKKKKKKHKSVALATSSDSASVTSKAKN---ALVDSESGAVREED 378
QY 143 -----ALSKHDEAIOR---HAAGAKRDEAFESSAFIOMQOQSLTSDLS 185
DB 379 VDRPAEAEQAQACSTEKHR-EAMORLEPTEHEESNSESASNAARHISEDRESDDVD 437
QY 186 IIAANPVYKFSNFISQGAATSLSDAKRAVDPIILYCOLVYKRELLVDLAI 245
DB 438 LQSA---VQOLREFIPDIDER---AAT---IRRYRDLDG- 469
QY 246 LYRKGAENHA-----SAVENANRYNKELAADTLFLKLLPEQALIGAVYHPIASSET 299
DB 470 LRKEFAAGCAIRFGKFSKAKE-----NKQLEKVVQDF-----SLTGIES 509
QY 300 SKAILNTYTFGVDPVPRIGN--RRYKFTNSYWNYSICSEAYMGNYMFGCSNVNPNP 357
DB 510 ADKLTLTDY---PEEKTLITNLRKHAF-----RLHICKGIARPKLV---Y 551
QY 358 IVSKMSD-----GFYTMNSDRRKLYITKHHQGW 387

DB 552 YRAKIFDVNNYKGRYNEEDTKLKAYHSLHGNDW 586
RESULT 5
ID 09X0M7 PRELIMINARY; PRT; 530 AA.
AC 09X0M7;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE METHYL-ACCEPTING CHEMOTAXIS PROTEIN.
GN TM1143.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
ON NCBI_Taxid=2336;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Uterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
DR EMBL: AE001771; AAD36219.1; -
DR HSSP: P02942; IQ07.
DR TIGR: TM1143; -
DR InterPro: IPR000122; Chemotaxis_transducer.
DR InterPro: IPR000727; T_SNARE.
DR Pfam: PF00015; MCPsignal; 1.
DR PRINTS: PR00260; CHEMTRNSDUCR.
DR SMART: SM00283; MA; 1.
KW Complete proteome.
SQ SEQUENCE 530 AA; 57928 MW; C0419A4FA6890CE CRC64;

Query Match 5.0%; Score 116; DB 2; Length 530;
Best Local Similarity 19.4%; Pred. No. 2.8;
Matches 54; Conservative 63; Mismatches 121; Indels 40; Gaps 9;

QY 15 ITSAKHGKRSVDVSLTKVET---ALKEASGSNEALEALEGKETOTKPDYRGATKI 71
DB 248 ITNQLGISKEMKNTSTRIESIASVOETTAGSEELISAKTNLADSAQOASFPDOSTOL 307
QY 72 LGSVGSALGKLNCGDATKIISGCLDIYAGIATTFGGPVGNGIGAVASFVSSILSLFTGSS 131
DB 308 AKRAGDALKKVI--EVTMTSNAKADVERVESF---QKGAELTISFETINAI----- 356
QY 132 AKNSVAVIDRALSKRDEAIORHAAGAKRDEAFESSAFIOMQOQSLTSDLSITIANV 191
DB 357 AEQTNLLAALNAALEAR-----AGEAGRGFVVADEIKRLAEESQOASENVRVYNEI 409
QY 192 -----PVYKFSNFISQGAATSLSDAK-----RAYVDPIILYCOLVYKRE 238
DB 410 RSIADKAGVSS---EITARVEGCTKLADKLSIYGAVERINEMLONIAALEEOT 466
QY 239 LLYVD--LAILYRKGAENHAHAVANANRYNKELAADT 273
DB 467 AAVDEITTAETAKNAEETITNSVKRVNARLQETISAST 504
RESULT 6
ID 09R9R9 PRELIMINARY; PRT; 306 AA.
AC 09R9R9;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)

Accession	Protein	Length
QY_16	TSKKGHGRSPVNSLLTRKVEFLAKEASGN	53
Db_2037	TEADRRKQTFYNNAYTAATLLDRTGSDNKNVYVQALQRTNTAKTALNDGARLMEAKN	2096
OY_54	LKGELQTK-----PDRVGATKTIILGSG-----SGLGLN-----SGDAT	88

```

Db 2097 TAKOOLAMSHLTNAOKANLLETOIERGTTVAGVGIQANAGLTNQMNLKOSIASKDAT 2156
QY KIISGCDIVAGIATTEGPGVMIGAV-----ASFVSSILSLFTG--- 129
Db 2157 KSSEDIYODANADLONAYDAYTNAEGTISATNNPEMNDPTINOKRASOVNASKSALNGDEK 2216
QY 130 -SASKNVAAYIDRALSKHDEAIORHAAKARDEFAESSAFIYVMKQOS----- 177
Db 2217 LAAKQKFAKSDIGRLTDLN---AQRTAANAEVQAPPLAATLAKKATSLNTAMNMLK 2273
QY 178 -----NLTDSD-----LSITAAVVPYKFSNFIOLESRIISQGAATTSLSDAK 220
Db 2274 HALAEKDNTRKSVNYTDDADQKQOAYDTAVTQAEATITNANGSNANETQVQAALNQLNQA 2333
QY 221 RAYDFILLICQVYVRELLVDLALIRKGAENHA--SAVENANRVKELAA-DTLDFL 277
Db 2334 NDLEKDNKRVAAQ--AKETAKRALASYSNLLNNAOSTAATSOIDNATTVADVTAQONTANEL 2390
QY 278 HKLIPEQALIGAVYHPISASETSKAILNYTKYFQVDPDRIGNRRYKFTSYNNTYSIC 337
Db 2391 N-----TAMQDONGINDONTVAKQVMT-----DADQ---GKDAYTNAVTNAQIL 2435
QY 338 SEA 340
Db 2436 DKA 2438

RESULT 9
Q9F5M7 PRELIMINARY; PRT; 788 AA.
AC Q9F5M7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE METHYL-ACCEPTING CHEMOTAXIS PROTEIN MCPX.
GN MCPX.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RA Muschler P.F.J., Scharf B., Schmitt R.;
RT "Methyl-accepting Chemotaxis protein MCPX of Sinorhizobium meliloti."
RL Submitted (Oct-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF312879; AAG34158.1;
DR InterPro: IPR000267; Asparagine_glutamase.
DR InterPro: IPR000122; Chemotaxis_transducer.
DR InterPro: IPR000658; DUF5.
DR InterPro: IPR003660; HAMF.
DR InterPro: IPR003661; His_KinA.
DR Pfam: PF00672; DUF5.1.
DR Pfam: PF00015; MCPsignal.1.
DR PRINTS: PR00260; CHEMTRNSDUCR.
DR SMART: SM00304; HAMF.2.
DR SMART: SM00388; HSKA.1.
DR SMART: SM00283; MA.1.
DR PROSITE: PS00144; ASN_GLN_ASE.1; UNKNOWN.1.
SQ SEQUENCE 788 AA; 83753 MW; 3681BC8667DE83F5 CRC64;

Query Match 4.88; Score 112; DB 2; Length 788;
Best Local Similarity 19.5%; Pred. No. 10;
Matches 59; Conservative 55; Mismatches 134; Indels 54; Gaps 9;
QY 15 ITSAGKGRSDVNSLTKVETALKEAGSGNPALEALGKGEIQTKPDRVGATKILGS 74
Db 518 LITGVNNAEAERAKVAASVEVASSDAQSGEVVQKAIAMGIDQSS---HEVSRIIGV 573
QY 75 VGSALGKINSGDATKIISGCDIVAGIATTEGPGVMIGAVASVSSILSLFTGSSAKN 134
Db 574 I-----DEIAFQTNILALNAGVEAARAGAGKGFVAVVAGVEVRELMO--RSANAAK 621

```

```

QY 135 SVAVIDRALSKHDEA-IORHAAKARDEFAESSAFIYVMKQOSNLNDSLSITIAANVPY 193
Db 622 EIKTLINTSAGQVREGVDLVKAGAGALEKIAEOVYQINGLRQJSSASQEA-----VGL 676
QY 194 YKFSNFIOLESRIISOG-----AATTSLSDAKRAYDFILLICQVYVRELLVDLAI 245
Db 677 KEINSAVNQMDQVQONAAWVEETTAASMALEDEARLALSALVARQIAPOAAQASAM 736
QY 246 LYRKGAENHAHAVENANRVKELAADTLDFLHKLIPEQALIGAVYHPISASETSKAILN 305
Db 737 L-RGTERMRRAAPAEHR-----PAQABRSAY-----SNSTORVIA 772
QY 306 YT 307
Db 773 KT 774

RESULT 10
Q9JBJ4 PRELIMINARY; PRT; 2026 AA.
AC Q9JBJ4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE PHAB PROTEIN.
GN PHAB.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=FAM18;
RX MEDLINE=20187481; PubMed=10722605;
RA Klee S.R., Nassif X., Kusecek B., Werker P., Beretti J.L., Achtman M.,
RA Tinsley C.R.;
RT "Molecular and biological analysis of eight genetic islands that
RT distinguish Neisseria meningitidis from the closely related pathogen
RT Neisseria gonorrhoeae."
RL Infect. Immun. 68:2082-2095(2000).
DR EMBL: AJ391284; CAB72074.1;
SQ SEQUENCE 2026 AA; 210271 MW; D04B9C1EB9C0085A CRC64;

Query Match 4.88; Score 112; DB 2; Length 2026;
Best Local Similarity 23.5%; Pred. No. 44;
Matches 80; Conservative 53; Mismatches 148; Indels 60; Gaps 16;
QY 26 VNSL-----LTKVERALKEAGSGNPALEALGK-LEGEIQTKPDRVGATKILGSVSGALG 80
Db 1639 VNSFQGEASAKIKTTFSDDYVAKQFA-HALAGCVSLVQGCCKD-----GALGAIVG 1689
QY 81 K-----LNSGDATKIISGCDIVAGIATTEGPGVMIGAV-ASFVSSILSL 125
Db 1690 EIVAEKSLGGRNPAITLDAEKHKVYSKIAGSVALNGDVTMANAAVAVYNNMLN 1749
QY 126 L-FTGSSAKNSVAAVIDR-ALSKHDEAIORHAAKARDEFAESSAFIYVMKQOSNLTDSD 183
Db 1750 FDSPTPAKHKQPKPDKTALKEKIQSIMPAAAGAMTNPDQKDAALWISNIRNGITG-- 1807
QY 184 LSITAAVVPYKFSNFIOLESRIISOGAATTSLSDAKRAYDFILLICQVYVRELLVDL 243
Db 1808 -PIVTSYGYVA-AGWAPPLIGTAGKKAISTCMNPSGCTVYVQAALAGAGIATGATV 1865
QY 244 AIIYRKQNA-EHVASAVENANRVKELAADTLDFLHKLIPEQALIGAVYHPIS-ASETSK 301
Db 1866 -----GNANEAAPGALSKAKKAKQALPTQTVKRLDGLDESKNIGAVNTPRINIANSTTR 1919
QY 302 AIIINYT--KTFQVP-----DVPRIIGNRRYFTNS 329
Db 1920 -----YTPMROTGPVSAGFEHVLGHHRPDIANNRSVFTIS 1956

```


DR HSSP; P07024; 20SH.
 DR InterPro: IPR002224; 5_nucleotidase.
 DR InterPro: IPR001064; Crystallin.
 DR InterPro: IPR000934; Ser_thr_phosphatase.
 DR Pfam; P001009; 5_nucleotidase; 2.
 DR PROSITE; PS00785; 5_NUCLEOTIDASE_1; UNKNOWN_1.
 DR PROSITE; PS00786; 5_NUCLEOTIDASE_2; 1.
 DR PROSITE; PS00225; CRYSTALLIN_BETGAMMA; UNKNOWN_1.
 FT NON_TER 865
 FT SEQUENCE 865 AA; 95050 MW; 70EB9AFA9CDF5665 CRC64;

Query Match 4.8%; Score 111; DB 2; Length 865;
 Best Local Similarity 22.6%; Pred. No. 14; Mismatches 178; Indels 138; Gaps 27;
 Matches 110; Conservative 61;

QY 14 AITSAKHGRSDVNSLLKRVETALKESGS---NEALEALEGLK--ETIQKPRDYGQ 67
 DB 280 AIISEPARAKYVKFNLTVE---KDESGTVLDKNADLLSLKGVQDPELTKLPPE 336
 QY 68 ATKILGSVGSALGKINSGLDQKTIISGLDIYAGT-----ATTEGCPVGM----- 111
 DB 337 AA--INDATAKIGELKGGDLAK---PDEVKGIPOSIVEDQGYTDFINEVOLYNSKKFL 389
 QY 112 --GIGAVASFSVSSILSLFT-----GSSAKNSVAAY--IDRALSKHDEALORHAGA 159
 DB 390 QTKGIDPRNNYVWSSAALFSPKALKEGPIKADVSNIYKFDNKL---YVKTNGKQL 444
 QY 160 KRDEASSAFIQVKQKQSNLSDSLT--IAANVPYKESNEIG--QLESRIQ--GAATTS 215
 DB 445 KKYMEENKFEFNKKY-----DGDLTISFDENVRYKYDMEGVNVEINAKDGERIEN 498
 QY 216 L---SDAKRAVDLILQLVYVARETLVLDALILYKGNAAHVASAVENANRVKELAAD 272
 DB 499 LKFSKDGKPEVSDVYVLYS-----NDYRNSGLAAG 530
 QY 273 TLDFLHKLIPQALIGAVYHPIASB--TSKAILNTKRYGVPVPRPIGR---RYKF 326
 DB 531 IMDSGE---HEKITYDVNDISAIKRLISDIYIN-----VKHCIVIRNVDGNMKI 577
 QY 327 TNSIYN-----TYSICSEAVMGNYMFGCSNVNPNIR-----VSKMSDGYFTMENS 374
 DB 578 TGNMNMNEQORALAVKLINE---GKIKLPTSNGRTPVKSVTWGEVGSFAAL--PEEKKE 633
 QY 375 RRKIYIKKHQGMGVLDEPQDGMHRIPLRHGKYMSSKMPNF-----MTMES 428
 DB 634 VEIPIILFND---FHGSLKESGNPAAKVF---GELKKVKEKNPVTIYVSGGDMYOGS 686
 QY 429 SASGYIR 435
 DB 687 ALSNMLK 693

RESULT 14
 Q9BJ59 PRELIMINARY; PRT; 1657 AA.
 AC Q9BJ59;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE RNA POLYMERASE II LARGEST SUBUNIT.
 GN RPAATL.
 OS Leptomonas seymouri.
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leptomonas.
 OX NCBI_TaxID=5684;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21169359; PubMed=11266558;
 RA Gilling G., Belliofatto V.,
 RT "Trypanosome spliced leader RNA genes contain the first identified RNA
 RL Nucleic Acids Res. 29:1556-1564(2001).
 DR EMBL; AF338253; AAK15346.1; -

SO SEQUENCE 1657 AA; 182912 MW; B9C42C166AF8AD6A CRC64;

Query Match 4.8%; Score 111; DB 5; Length 1657;
 Best Local Similarity 24.1%; Pred. No. 38; Mismatches 126; Indels 122; Gaps 20;
 Matches 95; Conservative 51;

QY 65 VQATKILGSVGSALGK-----NSGDATKIISGLDIYAGIATTFGCPVGM----- 111
 DB 592 VGATTK--GVGAAGPSLIHVIIFNERSGDEVAKFINGQRTITTYENYCFARSGVQDTYA 649
 QY 112 -----GIGAVASFSVSSILSLFTGSSAKNSVAAYIDRALSKHDEA 151
 DB 650 DASTLKEMNNVHLKTRQSVETIGAAAN--NGKLTAKMSLQSPEDAVNSALNKCREEA 707
 QY 152 IORHAGAKRDPABSAFIQVKQKQSNLSDSLTIAANVPYKESNIGOLE---SRI- 207
 DB 708 AKKALSNVR---TNSKVMIEAGSKGSDLNICQIAV-----FVGQNNVAGSRIP 754
 QY 208 -----SGAATTSLSDAKRAVDLILQLVYVARETL-----VD 242
 DB 755 FGFRRLPLPHMLDDYGETSRGMATRGVEGLQPEF---YFHTVAGREGLIDTAVKTS 811
 QY 243 LAIILYRK--GNAEHVASA---VENANRVNKEIADTLDFLHKLIPQALIGAVYHPIA 296
 DB 812 TGYLQRLIKRLLEDVHAAYDCTVRNANOELIQLA-----YGEDGLDGA-----R 855
 QY 297 SETSKAILNTKRYGVPDPV--RPIGNRRYKFTSNYMTYSICSEAVMGNYM--FRGCSN 352
 DB 856 IEGNQA-----PPIPHMANEMIDKRYEY-----NEESFSNNMGSGSYMDPVRSL 903
 QY 353 VRNPNIRYKSKSDGF--VTMENSDDRKLITYTKHQ 385
 DB 904 LRDPQ--SVSKLQEEFDQLVKRDRMSRLTIIDMEK 936

RESULT 15
 Q9WYR0 PRELIMINARY; PRT; 656 AA.
 AC Q9WYR0;
 DT 01-NOV-1999 (Tremblrel. 12, Created)
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE METHYL-ACCEPTING CHEMOTAXIS PROTEIN.
 GN TM0429.
 OS Thermotoga maritima.
 OC Bacteria; Thermotogales; Thermotoga.
 OX NCBI_TaxID=2336;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSB8 / DSM 3109;
 RX MEDLINE=99287316; PubMed=10360571;
 RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
 RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
 RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
 RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
 RA Heidelberg S.L., Smith H.O., Venter J.C., Fraser C.M.,
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from
 RL genome sequence of Thermotoga maritima."
 RT Nature 399:323-329(1999).
 RU Nature 399:323-329(1999).
 DR EMBL; AE001721; AAD3514.1; -
 DR HSSP; P02942; 1007.
 DR TIGR; TM0429;
 DR InterPro: IPR000122; Chemotaxis_transducer.
 DR InterPro: IPR000658; DUF5.
 DR InterPro: IPR003660; HAM.
 DR Pfam; PF00672; DUF5; 1.
 DR Pfam; PF00015; MCPSignal; 1.
 DR SMART; SM00304; HAM; 1.
 DR SMART; SM00283; HAM; 1.
 KW Complete proteome.
 SO SEQUENCE 656 AA; 72217 MW; EB932EC3505B143 CRC64;

Query Match 4.8%; Score 110.5; DB 2; Length 656;
 Best Local Similarity 20.1%; Pred. No. 9.9; Mismatches 107; Indels 95; Gaps 10;
 Matches 62; Conservative 45;

```

QY 24 SDVNSLTQVETALKEASGSENALEALEGL-KGEIQTKPDRYQATKILGSVGSALGR 81
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 393 NNMSSALTEVTSGVEEVAASAONISKITQDLTERSEAVTKAAREG-TERVEAVGVINK 450
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 82 LNSG-----DATKIISGCLDIVAGIA-----TTGCPYGMGIGAV 116
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 451 LKGSARQRDYLRELVDASAKTIGEIYDTISSIAEQTNLALNLAIEAARAGEAGRGFAV 510
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 117 ASFVSSITSLFTGSSAKNSVAVIDRALSKHDEAIQRHAAGAKRDFEASSAFIOVMKQ 176
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 511 ADEIRKL-----AESORATEDIAKMLSLRTTEHVENGSKEMFEGVDEIAVMGEE 562
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 177 SNLTSDSLIAANVPYKFSNFIGOLE--SRIQGAATTSLSDAKRAVDFILLYCQLV 233
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 563 -----VTKRFEILGRIEELINSMIENFAATAQEGGA----- 594
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 234 VARETLLVDLAILYRKGNAEHVASAVENANRUNKELADTLDFLHKLIPEQALIGAVYHP 293
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 595 -----AEEMASAMDNYTKI-VEGVESLNRMESLIEDQT----- 627
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 294 ISASETSKA 302
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 628 TSAKVSOA 636
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: May 8, 2002, 12:17:04
 Job time: 212 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 12:13:12; Search time 13.31 seconds
(without alignments)
1239.607 Million cell updates/sec

Title: US-09-647-522-5

Perfect score: 2322

Sequence: 1 MIKHLPMFLVLAITSAKH.....SGYIRSENNPQGHMSIT 450

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 100059 segs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	116.5	5.0	376	1 MDHC_YEAST	P22133 saccharomyc
2	112.5	4.8	964	1 PMPE_CHLTR	O64877 chlamydia t
3	110	4.7	1073	1 ACAG_ARATH	O91u41 arabidopsis
4	108.5	4.7	1023	1 HLVI_ECOLI	P09983 escherichia
5	108.5	4.7	1024	1 HLXA_ECOLI	P08715 escherichia
6	105.5	4.5	312	1 FAS3_RHOFA	P46375 rhodococcus
7	105	4.5	306	1 AER_ECOLI	P50466 escherichia
8	103	4.4	1456	1 RRPO_PVXCP	P22591 potato viru
9	102	4.4	393	1 FLAA_PSEAE	P21184 pseudomonas
10	102	4.4	1456	1 RRPO_PVXHB	O07630 potato viru
11	101.5	4.4	299	1 PYRB_ARCFU	O30130 atcheoglob
12	101	4.3	535	1 HTR1_HAUNI	P33741 halobacteri
13	100	4.3	437	1 YABE_BACSU	P37546 bacillus s
14	100	4.3	444	1 Y808_CHLPN	O92798 chlamydia p
15	99	4.3	948	1 PMAS_ARATH	O951b3 arabidopsis
16	99	4.3	1935	1 PYSS_CYCA	O90339 cyprinus ca
17	98.5	4.2	379	1 FLAA_VIRCH	O30858 vibrio chol
18	97.5	4.2	487	1 SR54_ENTHI	O15821 entamoeba h
19	97	4.2	825	1 GUN3_BACS4	P19570 bacillus sp
20	97	4.2	1934	1 MYH7_MESAU	P13540 mesocricetu
21	97	4.2	1935	1 MYH7_RAT	P02564 ratulus norv
22	95.5	4.1	3027	1 POLG_PYFVL	O05057 parsnip yel
23	95	4.1	381	1 YBD2_YEAST	P38199 saccharomyc
24	95	4.1	434	1 T147_HUMAN	O60664 homo sapien
25	95	4.1	483	1 KICL_MOUSE	O64291 mus musculu
26	95	4.1	675	1 ATKB_DEIRA	O92z20 deinococcus
27	95	4.1	1120	1 STFR_ECOLI	P76072 escherichia
28	95	4.1	1456	1 RRPO_PVX	P09395 potato viru
29	94.5	4.1	548	1 SR54_AOUAE	O67615 aquifex aeo
30	94.5	4.1	548	1 HLXB_VIRCH	P15442 vibrio chol
31	94.5	4.1	1706	1 CYAA_BORPE	P15318 bordetella
32	94	4.0	374	1 FLAI_BARBA	P35633 bartonella
33	94	4.0	450	1 ADFP_BOVIN	O9tun6 bos taurus

34	94	4.0	557	1 MCPS_ENTAE	P21822 enterobacte
35	94	4.0	1935	1 MYH7_HUMAN	P12883 homo sapien
36	93.5	4.0	427	1 HTSX_HAETN	P44001 haemophilus
37	93.5	4.0	580	1 IPAB_SHITL	P18011 shigella fl
38	93.5	4.0	903	1 MSP1_SCHPO	P87320 schizosacch
39	93.5	4.0	1807	1 TSC2_HUMAN	P49815 homo sapien
40	93	4.0	488	1 FLIC_PSEAE	P72151 pseudomonas
41	93	4.0	845	1 Y4FA_RHISN	P55439 rhizobium s
42	93	4.0	1295	1 BXA_CLOBO	P10845 clostridium
43	92.5	4.0	431	1 KICQ_HUMAN	O04695 homo sapien
44	92.5	4.0	553	1 MCP2_SALTY	P02941 salmoneilla
45	92.5	4.0	1790	1 LMBI_DROME	P11046 drosophilla

ALIGNMENTS

RESULT	ID	MDHC_YEAST	STANDARD	PRT	376 AA.
1	MDHC_YEAST				
AC	P22133:				
DT	01-AUG-1991 (Rel. 19, Created)				
DT	01-AUG-1991 (Rel. 19, Last sequence update)				
DT	30-MAY-2000 (Rel. 39, Last annotation update)				
DE	MALATE DEHYDROGENASE, CYTOPLASMIC (EC 1.1.1.37).				
GN	MDH2 OR YOL126C.				
OS	Saccharomyces cerevisiae (Baker's yeast).				
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;				
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.				
OX	NCBI_TaxID=4932;				
RP	[1]				
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 1-16.				
RX	MEDLINE=91094852; PubMed=1986231;				
RA	Minard K.I., McAlister-Henn L.;				
RT	"Isolation, nucleotide sequence analysis, and disruption of the MDH2				
RT	gene from Saccharomyces cerevisiae: evidence for three isozymes of				
RT	yeast malate dehydrogenase.";				
RL	Mol. Cell. Biol. 11:370-380(1991).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=S288C / FY1679;				
RX	MEDLINE=97051588; PubMed=8896265;				
RA	Casamayor A., Khalid H., Balcells L., Aldea M., Casas C., Herrero E.,				
RA	Ariño J.;				
RT	"Sequence analysis of a 13.4 kbp fragment from the left arm of				
RT	chromosome XV reveals a malate dehydrogenase gene, a putative Ser/Thr				
RT	protein kinase, the ribosomal L25 gene and four new open reading				
RT	frames.";				
RL	Yeast 12:1013-1020(1996).				
RN	[3]				
RP	SEQUENCE OF 1-34.				
RX	MEDLINE=87185517; PubMed=3552052;				
RA	Kopetzki E., Entian K.-D., Lottspeich F., Mecke D.;				
RT	"Purification procedure and N-terminal amino acid sequence of yeast				
RT	malate dehydrogenase isoenzymes.";				
RL	Biochim. Biophys. Acta 912:398-403(1987).				
CC	-1- FUNCTION: THE ISOENZYME MDH2 MAY FUNCTION PRIMARILY IN THE				
CC	GLYOXYLATE CYCLE.				
CC	-1- CATALYTIC ACTIVITY: L-MALATE + NAD(+) = OXALACETATE + NADH.				
CC	-1- SUBUNIT: HOMODIMER.				
CC	-1- SUBCELLULAR LOCATION: CYTOPLASMIC.				
CC	-1- INDUCTION: BY ACETATE AS CARBON SOURCE IN THE GROWTH MEDIUM.				
CC	-1- IS INACTIVATED BY ADDITION OF GLUCOSE (CARABOLITE INACTIVATION).				
CC	-1- MISCELLANEOUS: YEAST CONTAINS AT LEAST 3 MALATE DEHYDROGENASE				
CC	ISOENZYMES: A MITOCHONDRIAL (MDH1), A CYTOPLASMIC (MDH2) AND A				
CC	PEROXISOMAL (MDH3).				
CC	-1- SIMILARITY: BELONGS TO THE LDH FAMILY. MDH SUBFAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M62808: AAA34766.1; -
 DR EMBL: U41293: AAC49466.1; ALT_INIT.
 DR EMBL: U74868: CAA9145.1; ALT_INIT.
 DR PIR: S12937: DEBYMC.
 DR PIR: S05770: S05770.
 DR HSSP: P00346: 1MLD.
 DR SGD: S0005486: MDH2.
 DR InterPro: IPR001252: MDH.actsite.
 DR InterPro: IPR001236: 1dh.
 DR Pfam: PF00056: 1dh.
 DR PROSITE: PS00068: MDH.1.
 KM Oxidoreductase: Tricarboxylic acid cycle; NAD.
 FT INIT_MET 0
 FT ACT_SITE 181
 FT BINDING 184
 FT ACT_SITE 214
 SQ SEQUENCE 376 AA; 40599 MW; C08C63BF85E457EE CRC64;

Query Match 5.0%; Score 116.5; DB 1; Length 376;
 Best Local Similarity 20.2%; Pred. No. 0.54;
 Matches 72; Conservative 65; Mismatches 125; Indels 95; Gaps 17;

QY 13 LAITSAGHGRSDVNSLL-TKVEALKEASGS-----NEALELEGLKEGIOTKPD 63
 DB 15 IALIGAAGGGQSLTLKKAQLOQLKESKNSVTHILALYDNOEAINGVYADLSHIDT 74
 QY 64 RVGQATKILSVSGALKLSGDTAKITSG-----CLDIVGIAITTEGPGVMG 112
 DB 75 PIVSSSH--SPAGIEIENCLNNAIIVIPACVPKRGKTRDLEFVNNAGIISQDLSI--- 129
 QY 113 IGAVASVSSILSTFGSSAKNSYAVIDRALSKH---RDEAIORHAG-AKRDFAESSA 168
 DB 130 --AECCCLSFYFVLVINPNSILVPVVSNTLKNHPQSRSGIERRMGTGKLDIVRAST 187
 QY 169 FIVQMKQOSMLTD-----SPLSTIAAN-----VPYKFSNFIQGL----- 203
 DB 188 FLKRNIESGLTPRVNSMPDVYVGGHSGEITIFLFQSQNFSLANDQLKYLIRVOYG 247
 QY 204 ----ESRISGAATTSLSDA--KRAVDFTLLYQOLVMEETLLVLDLAILYRKGNAEHVA 256
 DB 248 GDEVYKAKNGKSGATLSMAHAGYKCVQFV-----SLL-----GNIDQIH 288
 QY 257 SA----VENNRVVKELAADTDLFLKLIPEQALIGAVYHPISSSESKAILNTKY 309
 DB 289 GTYYVPLKDNANPFIAGAD-----QLLP--LVDGADYFAIPLTITTKGV-STVDY 336

RESULT 2
 PMPE_CHLTR STANDARD: PRT: 964 AA.
 AC 084877;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE PROBABLE OUTER MEMBRANE PROTEIN PMPE PRECURSOR (POLYMORPHIC MEMBRANE
 DE PROTEIN E).
 GN PMPE OR CT869.
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID-813;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-D/UM-3/CX;
 RX MEDLINE-99000809; PubMed-9784136;
 RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
 RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
 RA Davis R.W.,
 RT "Genome sequence of an obligate intracellular pathogen of humans:
 RT Chlamydia trachomatis.";

RL Science 282:754-759(1998).
 CC -1- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
 CC (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: AE001360: AAC68467.1; -
 DR InterPro: IPR003368: DUF145.
 DR InterPro: IPR003357: OMP.
 DR Pfam: PF02415: DUF145.1.
 DR Pfam: PF02385: OMP.1.
 KM Outer membrane; Signal; Multigene family; Complete proteome.
 FT SIGNAL 1
 FT CHAIN 19 964
 SQ SEQUENCE 964 AA; 104703 MW; 1B998A7D2E571CE2 CRC64;

Query Match 4.8%; Score 112.5; DB 1; Length 964;
 Best Local Similarity 21.3%; Pred. No. 3.5;
 Matches 110; Conservative 61; Mismatches 184; Indels 161; Gaps 28;

QY 21 GKRSDV---NSLITVETALKEASGSNEALALEGLKEGIOTKPDVQATKILGSG 76
 DB 254 GNGNIVFYNNRCFKNVEFASSEAS-----DGAIKVTRLDVDTGNRGRIFFS- 301
 QY 77 SALGKLSGATKTIISGCDIVAGIATTPGPGVMGCAVASVSSILSTFGSSAKNSV 136
 DB 302 -----DNITKNYGAI---YAPVTLVDNGPTFYINNIAKNG 336
 QY 137 AAVIDRALSKRDEAIORHAG-----AGAKR-----DEASSAFIQV 172
 DB 337 GAIYIDGTSKSLISA-DRAHIIINENIVTNTANGSTSANPRRAIIVASSGELL 395
 QY 173 MKQOS-NLIDSD-LSITANVPYKFSNFIQLESRSOGAATTSLSDAKRAVD----- 224
 DB 396 GAGSQNLIFDYDIEVSNAGVSV-SFNKEADQYGSVFSGATVNSADFQRNLTQKTPAP 454
 QY 225 -----FILL--YCOLVVMKEL--LVLD-----AIL--YKGNAEHVAEVENNRVYKE 268
 DB 455 LITLSNGFLIEDHQAQLTVNRFOTGGVSLGNCAVLSCTYNGTGD-----SASNAITLKH 510
 QY 269 LAADTDLFLKLIPEQALIGAVYHPIASSETSAILNTY---KYGVPDVPRI---G 320
 DB 511 IGLN---LSSILKSGAEI-----PLTWVEPTNNSNNYTDATATFSLDVKLSLIDYG 561
 QY 321 NRRYKFTNSYNTWNTSICSEAYMGVYMGVYRGCSNVRNPIRVSKMSDGYTMENSDRRLYI 380
 DB 562 NSPYESTIDL--THALSSQ-----PWLSTISEASDNOLOSENIDFSGLVN 602
 QY 381 TKHD-OG--WGNG-FLDEPQDGHMR-----FPLP-RHGKVMV 414
 DB 603 PHIGWGLWTGNAKTODPEPASSAITTPDKANRRFRITLLTWLPAGVYPSFKHRSPLI 662
 QY 415 SSRKMPNMFYMESSASGYTRSWENNGPQGHWSIT 450
 DB 663 ANTLWGMMLLATESLKN---SALFLPSGHPFGIT 694

RESULT 3
 ACA9_ARATH STANDARD: PRT: 1073 AA.
 ID ACA9_ARATH
 AC 091041;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE POTENTIAL CALCIUM-TRANSPORTING ATPASE 9, PLASMA MEMBRANE-TYPE

DE (EC 3.8.3.8). (CA2+-ATPase, ISOFORM 9).
 GN AC9 OR AT3G21180 OR MSL8.3.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC Eustroids II; Brassicales; Brassicaceae; Arabidopsis.
 RN NCBI_TaxID=3702;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV COLUMBIA;
 RX MEDLINE=20277480; PubMed=10819329;
 RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
 RT features of the regions of 4,504,864 bp covered by sixty pl and YAC
 RT clones.";
 RL DNA Res. 7:131-135(2000).
 CC -1- FUNCTION: THIS MAGNESIUM DEPENDENT ENZYME CATALYZES THE HYDROLYSIS
 CC OF ATP COUPLED WITH THE TRANSLLOCATION OF CALCIUM FROM THE CYTOSOL
 CC OUT OF THE CELL OR INTO ORGANELLES (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + CA(2+)(CIS) = ADP + PHOSPHATE +
 CC CA(2+)(TRANS)
 CC -1- ENZYME REGULATION: ACTIVATED BY CALMODULIN (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- DOMAIN: THE N-TERMINUS CONTAINS AN AUTOINHIBITORY CALMODULIN-
 CC BINDING DOMAIN, WHICH BINDS CALMODULIN IN A CALCIUM-DEPENDENT
 CC FASHION (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
 CC (E1-E2 ATPASES). SUBFAMILY IIB.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AB023045; BAB01709.1;
 DR InterPro: IPR001757; E1-E2-ATPase.
 DR InterPro: IPR001454; Hydrolyase.
 DR InterPro: IPR000661; Na,K-ATPase.
 DR Pfam: PF00122; E1-E2-ATPase; 1.
 DR Pfam: PF00702; Hydrolyase; 1.
 DR PRINTS: PR00119; CATTATPASE.
 DR PROSITE: PS00154; ATPASE_E1_E2; 1.
 KW Hydrolyase; Calcium transport; Transmembrane; Phosphorylation;
 KW ATP-binding; Metal-binding; Magnesium; Calmodulin-binding;
 KW Multigene family; Hypothetical protein.
 FT DOMAIN 1 181
 FT TRANSMEM 182 202
 FT DOMAIN 203 220
 FT TRANSMEM 221 241
 FT TRANSMEM 242 369
 FT TRANSMEM 370 389
 FT TRANSMEM 390 426
 FT TRANSMEM 427 444
 FT DOMAIN 445 844
 FT TRANSMEM 845 863
 FT TRANSMEM 864 874
 FT TRANSMEM 875 895
 FT TRANSMEM 896 915
 FT TRANSMEM 916 938
 FT TRANSMEM 939 950
 FT TRANSMEM 951 975
 FT TRANSMEM 976 993
 FT TRANSMEM 994 1015
 FT TRANSMEM 1016 1025
 FT TRANSMEM 1026 1047
 FT TRANSMEM 1048 1073
 FT DOMAIN 1048 1073
 FT MOD_RES 44 35
 FT MOD_RES 482 482
 FT MOD_RES 789 789
 FT METAL 793 793

SO SEQUENCE 1073 AA; 117488 MW; 5AF6220EBD908069 CRC64;
 Query Match 4.7%; Score 110; DB 1; Length 1073;
 Best Local Similarity 22.8%; Pred. No. 5.9;
 Matches 93; Conservative 51; Mismatches 118; Indels 146; Gaps 20;
 QY 9 LFVLAITSAAKGRSD-----VNSLITKVTALKEAGSGNEALELGLK 55
 DB 197 ILIIAAVTSIALGIGIKTEGLKGMIDGGSIAFAVLIVYVSDYRSIQ--FQNLDEK 254
 QY 56 GEIOTKPDVGOATKI-----LGSVSAIGKINSQDPATKI----- 90
 DB 255 RNIDLEVRGGRVTKISTIDVYVGVDPRLRIGDDVPADGVLSIGSLAIDESSMTGSEKI 314
 QY 91 -----ISGC-----LDIVAGIATVTEG-----GPGMGIGAVAS 118
 DB 315 VHKDQSPFLMCGKVDGVGNMLVTGVINTGMLVASISEDPTEGPLOVRNLGLAT 374
 QY 119 FVSSILFLTGSSAKNSVAVIDALSKHDEALORHAAKAKRPAESSAFTQWKQOSN 178
 DB 375 FI-GIVGL-----SVALVIVLAL-----LVRYFTGTDTGNGATOFIKGTTSSID 418
 QY 179 LTDS-----DLSTIAANVP-----VYKFSNFICQ--LESRIIS--OGAATT 214
 DB 419 IVDQVAFITFAVTVVAVVAVVEGELPLAVTLTLAISMKRMADKALVRLSACETMGSAFT 478
 QY 215 SLSDAKRAVDIFLLYCOLVYVRETLVLDLALYRKGAHVAASAVENARVKEADTL 274
 DB 479 ICSDKTGT-----LTLMQMTVV-ET-----VAGSKMDVA--DNPGLPRLVA--- 519
 QY 275 DFLHKLPE---QALIGAVVHP-----ISASETSALINTYTFYGV 312
 DB 520 -----LISEGYAOMTNTGNIHFPHKDGVEIISGSETERALISWAYKLGM 562
 RESULT 4
 HLVI_ECOLI STANDARD; PRT; 1023 AA.
 AC P09983;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE HEMOLYSIN, CHROMOSOMAL.
 GN HLVA.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OC NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-J96 / SEROTYPE O4;
 RX MEDLINE=85234404; PubMed=3891743;
 RA Felmlee T., Pellett S., Welch R.A.;
 RT "Nucleotide sequence of an Escherichia coli chromosomal hemolysin.";
 RL J. Bacteriol. 163:94-105(1985).
 RN [2]
 RP SEQUENCE OF 1-44 FROM N.A.
 RC STRAIN-2001;
 RX MEDLINE=85258115; PubMed=3894051;
 RA Nicaud J.-M., Mackman N., Gray L., Holland I.B.;
 RT "Characterisation of HlyC and mechanism of activation and secretion
 RT of haemolysin from E. coli 2001.";
 RL FEBS Lett. 187:339-344(1985).
 CC -1- FUNCTION: BACTERIAL HEMOLYSINS ARE EXOTOXINS THAT ATTACK BLOOD
 CC CELL MEMBRANES AND CAUSE CELL RUPTURE BY MECHANISMS NOT CLEARLY
 CC DEFINED.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING
 CC CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC
 CC ACTIVITY.
 CC -1- DOMAIN: THE THREE TRANSMEMBRANE DOMAINS ARE BELIEVED TO BE
 CC INVOLVED IN PORE FORMATION BY THE CYTOTOXIN.

CC or send an email to license@lsb-sib.ch.

CC EMBL: M14107; AAA98233.1; -

DR InterPro: IPR001343; Hemlysn_Ca_bind.

DR InterPro: IPR003353; RTX_N.

DR Pfam: PF00333; hemolysinCabin; 2.

DR PRINTS: PR00382; RTX; 1.

DR PROSITE: PS00330; CABNNGRPT.

KW Hemolysis; Toxin; Cytolysis; Cytotoxin; Repeat; Calcium; Transmembrane; Lipoprotein; Palmitate; Plasmid.

FT TRANSMEM 238 260 POTENTIAL.

FT TRANSMEM 268 327 POTENTIAL.

FT TRANSMEM 365 411 POTENTIAL.

FT DOMAIN 724 870 16 x REPEATS, GLY-RICH.

FT REPEAT 724 729 1.

FT REPEAT 733 738 2.

FT REPEAT 742 747 3.

FT REPEAT 751 756 4.

FT REPEAT 760 765 5.

FT REPEAT 769 774 6.

FT REPEAT 778 783 7.

FT REPEAT 787 792 8.

FT REPEAT 796 801 9.

FT REPEAT 807 812 10.

FT REPEAT 817 822 11.

FT REPEAT 826 831 12.

FT REPEAT 835 840 13.

FT REPEAT 844 849 14.

FT REPEAT 856 861 15.

FT REPEAT 865 870 16.

FT LIPID 564 564 PALMITATE.

FT LIPID 690 690 PALMITATE.

SEQUENCE 1024 AA; 110201 MW; 83944917F76C945B CRC64;

Query Match 4.7%; Score 108.5; DB 1; Length 1024;
Best Local Similarity 20.7%; Pred. No. 7.1; Indels 73; Gaps 12;
Matches 69; Conservative 52; Mismatches 139;

QY 24 SDVNSLTETALKEAGSGNEALELGEIKGEIQTP--DRYGOATKILGVSALGK 81
DB 200 NNVNSFSQQLNTL-----GSVLNTRKHLNGCVKMLQNLPLMDINGAGLDIVSGILSLISA 254
QY 82 ---LNSGDA---TKIISGCLDIYAGTTFGCGYGMKIGAVASFVSSILSLFTGSSAKNS 135
DB 255 SFLISNADADTRTKAAG---VELTTKVLGVNGKGISQYIIAQRAAQGLSTSAAGL 309
QY 136 VAAVIDRALSK-----HRDEAIORHAGAKRDPFASSAFIQVMKQOSNLTDLS 165
DB 310 IASAVTLAISPLSFLSIADKFKRANKLEIYSQRFKRLGIDGDSULAFHKEGAIDASLT 369
QY 186 IIAANPVYKFSNFIGLESRIISOGAATTSLSDAKRAVDLILLYCOLVYMRETLLVDLAI 245
DB 370 TI-----STVLASVSSGIS-AAATITLVGAPVSA-----LVGAVTGIIISILE 411
QY 246 LYKKGNAEHVASAV-----ENANRVKLEADTLDELHLKILPEQALIGAVYHPISASETS 300
DB 412 ASKQAMFEHVASKMAVDVIAEMEKKHGNYFENGVDARHAAFL-----DNF 457
QY 301 KAILNTKYGVPDPRPIGNRRYKFTNSWYNT 333
DB 458 KILSYNKEYSV-----ERSVLITQOHMT 482

RESULT 6
ID PAS3.RHOFA STANDARD; PRT: 312 AA.
AC P46375;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update).
DE HYPOTHETICAL 33.6 KDA PROTEIN IN FASCINATION LOCUS PRECURSOR (ORF3).

GN FAS3.
OS Rhodococcus fascians.
OG Plasmid pFID188.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Nocardiaceae; Rhodococcus.
OX NCBI_Taxid=1828;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DI188;
RA MEDLINE=94222824; PubMed=8169198;
RA Crespi M., Vereecke D., Temmerman W., van Montagu M., Desomer J.;
RT "The fas operon of Rhodococcus fascians encodes new genes required
for efficient fasciation of host plants."
RL J. Bacteriol. 176:2492-2501(1994).
CC -1- FUNCTION: THE FAS-OPERON ENCODES GENES INVOLVED IN CYTOKININ
PRODUCTION AND IN HOST PLANT FASCINATION (LEAFY GALL).
CC -1- INDUCTION: DURING THE INTERACTION WITH HOST PLANTS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.lsb-sib.ch/announce/>
or send an email to license@lsb-sib.ch).

CC EMBL: Z29635; CAA82743.1; -
KW Hypothetical protein; Plasmid; Signal.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 312 HYPOTHETICAL 33.6 KDA PROTEIN IN
FT FT FASCINATION LOCUS.
SEQUENCE 312 AA; 33574 MW; 849F168ABD4404F CRC64;

Query Match 4.5%; Score 105.5; DB 1; Length 312;
Best Local Similarity 22.2%; Pred. No. 2.4;
Matches 70; Conservative 48; Mismatches 93; Indels 105; Gaps 16;

QY 54 LKEIOTKPRVGOATKILGVSALGK-----LNSGATKILSGCLDIYAGIAT 103
DB 25 MSGALETDP-----RAVYVLADIGAHLEFKAAIADPNRVINVGIREQLMVG---VAGGLAM 77
QY 104 TFGPGVGMGAVASV--VSSILSLFTGSSAKNSVAAIYDIALSHRDAIQRHAAKRD 162
DB 78 CGMRPV---VHTVAALFVERPLEQIKLNFADQDQAVL-----VSMGASYD 120
QY 163 FASSAFIQVMKQOSNLTDLSII-----AANVP-----VYKFS 197
DB 121 LSE-PAF-----SHPTEDITVIDSMRWYVHVPHQPEADLLESPLGGRVTL-- 170
QY 198 NFIGLESRIIS-----QGAATTSLSDAKRAVDLILLYCOLVYMRETLLVDLAIY-- 247
DB 171 ---RLSSQVNRYPHVRGTSFPIKYGTRGVVLAVPCDAVLASVMDITLVYATI 226
QY 248 RKNAGHVASAVNANRVKLEADTLDELHLKILPEQALIGAVYHPISASETS--KAILN 305.
DB 227 RPPDATGGLCAVQAVNRPVVLV-----EPYLAGTSAHOVSLSVSHPHRLS 274
QY 306 Y-----TKYFGVPD 314
DB 275 LGVREMEDRHVGTDP 290

RESULT 7
ID AER.ECOLI STANDARD; PRT: 506 AA.
AC P50466;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE AEROTAXIS RECEPTOR.
GN AER OR AIR OR B3072.
OS Escherichia coli.

CC Bacteriophila; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CC Escherichia.
OX NCBI_TaxId=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.A., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RL "The complete genome sequence of Escherichia coli K-12";
RL science 277:1453-1474(1997).
RN [2]
RP MEDLINE=97334134; PubMed=9190831;
RX Bibikov S.I., Biran R., Rudd K.E., Parkinson J.S.;
RL "A signal transducer for aerotaxis in Escherichia coli.";
RL J. Bacteriol. 179:4075-4079(1997).
RN [3]
RP FUNCTION.
RX MEDLINE=98021405; PubMed=9380671;
RA Rebhargada A., Johnson M.S., Harding C.P., Zuccarelli A.J.,
RA Fletcher H.M., Zhulin I.B., Taylor B.L.;
RL "The aer protein and the serine chemoreceptor Tsr independently sense
RT intracellular energy levels and transduce oxygen, redox, and energy
RT signals for Escherichia coli behavior";
RL Proc. Natl. Acad. Sci. U.S.A. 94:10541-10546(1997).
CC -1- FUNCTION: SIGNAL TRANSDUCER FOR AEROTAXIS. THE AEROTACTIC
CC RESPONSES IS THE ACCUMULATION OF CELLS AROUND AIR BUBBLES. THE
CC NATURE OF THE SENSORY STIMULUS DETECTED BY THIS PROTEIN IS THE
CC PROTON MOTIVE FORCE OR CELLULAR REDOX STATE. IT USES A FAD
CC PROSTHETIC GROUP AS A REDOX SENSOR TO MONITOR OXYGEN LEVELS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE. INNER MEMBRANE
CC (POTENTIAL).
CC -1- SIMILARITY: IN THE N-TERMINAL TO NIPL AND FIXL AND OTHER PROTEINS
CC IMPLICATED IN OXYGEN-SENSING.
CC -1- SIMILARITY: BELONGS TO THE CHEMOTAXIS SENSORY TRANSDUCERS FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as their content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
CC or send an email to license@isb-sdb.ch).
CC
CC - EMBL; U28379; AAA89151.1; -
CC EMBL; AE000389; AAC76107.1; -
CC Ecogene; EG12955; aer.
DR InterPro: IPR0001022; Chemotaxis_transducer.
DR InterPro: IPR0000658; DUF5.
DR InterPro: IPR001610; PAC.
DR InterPro: IPR000014; PAS.
DR Pfam: PF000672; DUF5; 1.
DR Pfam: PF00015; MCPsignal; 1.
DR Pfam: PF00785; PAC; 1.
DR SMART: SM00283; MA; 1.
DR SMART: SM00086; PAC; 1.
DR SMART: SM00091; PAS; 1.
KW Chemotaxis; Transducer; Transmembrane; Methylation; Periplasmic;
KW Inner membrane; FAD; Flavoprotein; Complete proteome.
FT DOMAIN 1 166
FT TRANSEM 167 186
FT DOMAIN 187 190
FT TRANSEM 191 209
FT DOMAIN 210 506
FT POTENTIAL.
FT PERIPLASMIC (POTENTIAL).
FT CYTOPLASMIC (POTENTIAL).
FT POTENTIAL.
FT CYTOPLASMIC (POTENTIAL).
FT DOMAIN 506 AA; 55065 MW; 913DECF14EE3FD08 CXC64;
SQ SEQUENCE

```

0Y 26 VNSLTFWEETALKASQSNSE-----ALAEALGELGELIOTRPDRYQATFKILIGSGA 78
Db 299 MNQMAASVKQNSATASASADKISTITASNAAYOGGEMATTVKTHDD-1ADSTQKIGTITSL 357
QY 79 LGLNSGDATKIIISGCLDIYAGIATTFEGPGVMGICGAVASFSVSLISLFTGSSAKNSVAA 138
Db 358 INDI-----AFQYNIITALNAAVEAARAGEQKGFAYVAGEVTHLAS--RSANAANDIRK 409
QY 139 VIDBALSKHREALORHAAGAKKRPFASSAFIOWMKQSNLTJDSLSIIAANPYVKFSN 198
Db 410 LIDNSADRVQSGQCVHAGATKME-----DIVAQQKNVT----- 443
QY 199 FIGUESRISQGA--ATTSLSDAKRAVDITLLYCO--LVVMRETLVDLAILRKGAH 254
Db 444 ---QLIAQISHTLEQADGLSILTRAVDEINLTQKNAELVEBSAQVSAMV-----KH 493
QY 255 VASAVEANA 262
Db 494 RASRLEDA 501

```

	RESULT	8			
	RPO_PVXCP				
ID	RRPO_PVXCP	STANDARD:	PRT:	1456 AA.	
AC	P22591;				
DT	01-AUG-1991 (Rel. 19, Created)				
DT	01-AUG-1991 (Rel. 19, Last sequence update)				
DT	20-AUG-2001 (Rel. 40, Last annotation update)				
DE	RNA REPLICATION PROTEIN (165 KDA PROTEIN) (ORF 1) [CONTAINS: RNA-				
DE	DIRECTED RNA POLYMERASE (EC 2.7.7.48); PROBABLE HELICASE].				
OS	Potato virus X (strain CP) (PVX).				
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Potexvirus.				
OX	NCBI_TaxID=12184;				
RN	[1]				
RP	SEQUENCE FROM N.A..				
RX	MEDLINE=90364772; PubMed=2392880;				
RA	Orman B.E., Celnik R.M., Mandel A.M., Torres H.N., Mentaberry A.N.;				
RJ	"Complete cDNA sequence of a South American isolate of potato virus				
RT	X";				
RL	Virus Res. 16:293-306(1990).				
CC	-I- FUNCTION: RNA-REPLICATION. THE CENTRAL PART OF THIS PROTEIN				
CC	POSSIBLY FUNCTIONS AS A ATP-BINDING HELICASE.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	The European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
DR	EMBL; M31541; AAA47178.1; -.				
DR	EMBL; X55802; CA33924.1; -.				
DR	PIR; S14005; S14005.				
DR	InterPro; IPR000606; Viral_helicase1.				
DR	Pfam; PF01443; Viral_helicase1.				
KW	ATP-binding; Helicase; RNA replication; RNA-directed RNA polymerase;				
KW	Transferase.				
FT	NP_BIND 735 742 POTENTIAL.				
SQ	SEQUENCE 1456 AA; 165301 MW; 489BA57EA070BD2E CRC64;				
QY	Query Match 4.4%; Score 103; DB 1; Length 1456;				
	Best Local Similarity 18.7%; Pred. No. 27;				
	Matches 104; Conservative 70; Mismatches 187; Indels 194; Gaps 26;				
QY	52 EGLKELIQTKPDYVGQAATRIISVGS-----ALGKLNSG-DATKIISG---CLDIYA 99				
	I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :				
Db	714 EYLTKETISLPDKRYAACVHIGAGSGSKSHAIKALREIKGSDITVVLPNTNELRDMSK 773				
QY	100 GIATT-----FGG-----PYGMGIGAVASVSST-ISLPTGSSAK 133				
	: I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :				

```

Db 774 KVPNTEPYMFYKALIGTGSIYIFDDYSKLPGY-IALVSPSTIKILLINGDSRQ 832
QY 134 NSVAVIDRALSKHDEAIQ-----RHAAGA-----KDFPSSAFIOYMAQOQSLTSD 183
Db 833 SVYHETSDASIRHLGPATEVFAKYRYLNATHRNKDLAN--MLGVSERTGTTEIS 889
QY 184 LSI-----IAANVP-----VYKFSNFIGQLSRI-----SQGA 212
Db 890 MSSEPLEVPLVPSDEKRRRLMTGRNDTFTYACCGLTFRKQIVLDHNTQVCSANVM 949
QY 213 TTSLSDAKRAVDFT-----LTYCOLVYMETLLVDLAIYRKGNAEHV 255
Db 950 YSALSRAVDRIHFIETANSASFWEKLDSTPYLKFSLVVEHAKLKE-----YEPREAPFI 1005
QY 256 AS-----AYENANRVKELADTLD-----FLHKLIP 282
Db 1006 KEPEPOTMCVENESEVLEEKKELEKFDREIHSDAHGSNCVOTEDTTIQLFSHQAK 1065
QY 283 EQLIGAVYHPIASSETS-----LNYTKYFGVDPVPRIGNRKRYKT 327
Db 1066 DETLMATIDARKLITSNOESNREFLSKRDIGDVLFLMYQKAMGLPKPEPI-----FS 1118
QY 328 NSWNNTYSICSEAYMGNYMFRGCSNVNRPNIIVRSKMSDGFYTMENSDRKLITKHDGM 387
Db 1119 QEVWEA---CAHEVQSKYLSKSCNLLNGTVROS-----PDPDENKIMVFLKSQ-- 1164
QY 388 GWGLDEDPDQGHMRFIP-LRHGKRYV-----SSKRPMNEMWYESSASGYIR 435
Db 1165 -WVTKVEKLG-----LPKIKPGQTIAFYQOTVLMFGTWARVMRQAFQPEKVEPI- 1215
QY 436 SWENNPQOCHWSIT 450
Db 1216 NCETTPEDMSAMALS 1230

```

```

RESULT 9
FLAA_PSEAE
ID FLAA_PSEAE STANDARD; PRT; 393 AA.
AC P21184;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE FLAGELLIN TYPE A.
GN FLAA.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A. AND SEQUENCE OF 1-5.
RC STRAIN=PAK;
RX MEDLINE=91072275; PubMed=2123866;
RA Totten P.A., Iori S.;
RT "Characterization of the type a flagellin gene from Pseudomonas
RT aeruginosa PAK.";
RL J. Bacteriol. 172:7188-7199(1990).
CC -I- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO
CC FORM THE FILAMENTS OF BACTERIAL FLAGELLA.
CC -I- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M57501; AAA63458.1; -
DR PIR; A37853; A37853.
DR InterPro; IPR001029; Flagellin_C.
DR InterPro; IPR001492; Flagellin_N.
DR Pfam; PF00700; Flagellin_C; 1.

```

```

DR Pfam; PF00669; Flagellin_N; 1.
DR PRINTS; PR00207; FLAGELLIN.
DR ProDom; PD000316; Flagellin_C; 1.
KW Flagella.
FT INT MET
SQ SEQUENCE 393 AA; 39909 MW; 828B551543F73C29 CRC64;

```

```

Query Match 4.48; Score 102; DB 1; Length 393;
Best Local Similarity 21.38; Pred. No. 5.7;
Matches 71; Conservative 39; Mismatches 92; Indels 132; Gaps 13;

```

```

QY 11 IVALITSKHKRBDVNSLTKVETALKKAGSN-EAALALEGLKGIOTKPPRVQAT 69
Db 71 ISLAQT-AEGALQOSTNIIQMRDLISQANGSNSDSERTALNEAKLOKELDRISWT 129
QY 70 ----KIL-GSVGSALGLKNGDATTIT-----SG 93
Db 130 TFGSKRLIDGSFQVYASQVQVGA-ANETISVIDEMSAESLNGTYFKADGGAVTAATASG 188
QY 94 CLDIYAGIATTFGGPVGM----- 111
Db 189 TVDIAIGI--TGSAAVNVKVDKMGNETAEQAARKIAAVNDANYGIGAFSDGDTISYVS 246
QY 112 ----GIGAVASFVSLSLFTGSSA-----KNSVAVID----RALSKH 147
Db 247 AGKDGSAITSVASGVYADIGSTVGTAAGVAPSAFAKTNQTVAKIDISTAKLSRR 306
QY 148 ---RDEAIORHAQAKRDEAFSSAFIYMKQOQN-----LTMSDLSTIAANVPV 193
Db 307 AGDRTTAIKQIDASVPVSVAVQNFNDTINLKNIGENVSAARGRIEDTFAETANL-- 364
QY 194 YKFSNFIGLESRTISOGAATTSLSDAKRAVDFTL 227
Db 365 -----TKNOYLQQAQGTAILAQAQDLPSVL 389

```

```

RESULT 10
RRPO_PVXHB
ID RRPO_PVXHB STANDARD; PRT; 1456 AA.
AC Q07630;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE RNA REPLICATION PROTEIN (165 KDA PROTEIN) (ORF 1) [CONTAINS: RNA-
DE DIRECTED RNA POLYMERASE (EC 2.7.7.48); PROBABLE HELICASE].
OS Potato virus X (strain HB) (PVX).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potexvirus.
OX NCBI_TaxID=73488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94015010; PubMed=8409947;
RA Querol M., van der Vlugt R., Goldbach R., Salazar L.F.;
RT "RNA sequence of potato virus x strain HB.";
RL J. Gen. Virol. 74:2251-2255(1993).
CC -I- FUNCTION: RNA-REPLICATION. THE CENTRAL PART OF THIS PROTEIN
CC POSSIBLY FUNCTIONS AS A ATP-BINDING HELICASE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X72214; CA51012.1; -
DR InterPro; IPR000606; Viral_helicase1.
DR Pfam; PF01443; Viral_helicase1; 1.
KW ATP-binding; Helicase; RNA replication; RNA-directed RNA polymerase;
KW Transferrase.
FT NP_BIND 735 742 ATP (POTENTIAL).
SQ SEQUENCE 1456 AA; 165138 MW; 10730A1EDDAFAFBA CRC64;

```


Query Match 4.4%; Score 102; DB 1; Length 1456;
 Best Local Similarity 18.7%; Pred. No. 31; Indexes 194; Gaps 26;
 Matches 104; Conservative 70; Mismatches 187;

QY 52 EGLKEIQTDRVGVGATKILGVS-----ALGKLSNG-DATKILISG---CLDIVA 99
 DB 714 EYLKGIIESPERKVAACVIGHAGSGSKSHALQALREIGSGIDITVLPINELRLDMSK 773
 QY 100 GIAT-----FGC-----PYGMGIGAVSFVSI-LSLFTSSAK 133
 DB 774 KVPTEPEYMERKYEKALIGGSGIVIPDDYSKLPFGY-IEALVFSFKIKILITGDSRQ 832
 QY 134 NSVAVIDRALSKRDBAIO-----RHAAGA-----KRDAESSAFLQWKKQSNLSDSD 183
 DB 833 SVHETSDDSIRHLGPAITEVFAKCYRYLNATHRKKNLAN--MLGVSEKGTETIS 889
 QY 184 LSI-----IAANP-----YKFSNFIGLESRI-----SOGAA 212
 DB 890 MSSEFEGVPTLVPSDEKRLKMGCTGRNDFTYAGCGGLKPKVOYLHDHNTQVCANVM 949
 QY 213 TTSISDAKRAVDI-----LLYQLVNARETLLVLAIIKRGNAEHV 255
 DB 950 YTALSRAITDRHFTINTSANSAPFEKLDSPYLKTFLSVREQALRE---YEFVEAPEPI 1005
 QY 256 AS-----AVENANRYNKELAADTLD-----FLHKLIP 282
 DB 1006 REPEPQHMCVENESLEEKELKEKREHSEAHGNSCQVOTDTTVQVLFSHQAK 1065
 QY 283 EQALIGAVYHPISASETSKA-----ILNTRYEGVPDVPPIGNRRYKFT 327
 DB 1066 DETLMTATIDARLKTSGNESNFRFELSKRIDGVLFLNYGRANGLPREPI-----FS 1118
 QY 328 NSWNTYSISSEAYMGVYMRGCSNVNPNIRVSKMSDGYTWENSORRKLITTKHQGW 387
 DB 1119 QEWEEA---CAHEVOSKYLSKSKCNCLNGTVROS-----PDFENKIMVFLKSO- 1164
 QY 388 GWGLDEDDPGDQGHMRFIP-LRHGKYVW-----SSKRPMNFMNMESASXYIR 435
 DB 1165 -WTKYVKLG-----LPKIKPGQITIAAFYQOTVLMFGTAMRYMFRQAFQPKVEYFI- 1215
 QY 436 SWENNPQPGHWSIT 450
 DB 1216 NCERTPEDMSAMALS 1230

RESULT 11
 PYRB_ARCFU STANDARD; PRT; 299 AA.
 ID AC 030130;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE ASPARTATE CARBAMOYLTRANSFERASE (EC 2.1.3.2) (ASPARTATE TRANSFERASE) (ATCASE).
 GN PYRB OR AF0106.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
 OC Archaeoglobus.
 NCBI_TaxID=2234;
 RX SROUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Klein H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Retchum K.A., Dodson R.J., Hickey E.K., Peterson J.D.,
 RA Richardson S., Kellavag A.R., Graham D.E., Kyriades N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.R., Badger J.H., Glodex A., Zhou L.,
 RA Overbeek R., Cocayne J.D., Weidman J.F., McDonald L., Utterback T.,
 RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,

RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 reducing archaeon Archaeoglobus fulgidus.";
 RL Nature 390:364-370(1997).
 CC -1- CATALYTIC ACTIVITY: CARBAMOYL-PHOSPHATE + ASPARTATE =
 CC ORTHOPHOSPHATE + N-CARBAMOYLASPARATE.
 CC -1- SUBUNIT: SECOND STEP IN PYRIMIDINE BIOSYNTHESIS.
 CC -1- SIMILARITY: BELONGS TO THE ATCASES/OTCASES FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AE001099; AAB91125.1; -
 DR TIGR: AF0106;
 DR InterPro: IPR002029; Carbmyltransf_asor.
 DR Pfam: PF00185; OTCase; 1.
 DR Pfam: PF02729; OTCase_N; 1.
 DR PRINTS: PR00100; ATCASE.
 DR PRINTS: PR00101; ATCASE.
 DR PROSITE: PS00097; CARBAMOYLTRANSFERASE; 1.
 DR Pyrimidine biosynthesis; Transferase; Complete proteome.
 SQ SEQUENCE 299 AA; 33633 MW; 70F2FC2EBACDD720 CRC64;

Query Match 4.4%; Score 101.5; DB 1; Length 299;
 Best Local Similarity 22.6%; Pred. No. 4.3; Indexes 97; Gaps 16;
 Matches 75; Conservative 38; Mismatches 122;

QY 23 RSDVNSLITVEFETALKEASGNEALELEG-LKEGETQTPDVGATKILGVSALGK 81
 DB 13 RSDVNSLITVEFETALKEASGNEALELEG-LKEGETQTPDVGATKILGVSALGK 81
 QY 82 LNSGDAITKIISSGLDIVAGIATTFGGVGVGIGAVASVSSILSLTGSAAKNSVAVID 141
 DB 66 L-----GGDV-----IMWTAQEAASI-----AKGETLADTI 91
 QY 142 RALSKHDEAIOHAGAKRDEFASSAFIOYMK-----QSNLTSDLSIIANVYVK 195
 DB 92 RIVSGYCDALIVIRPLEGAARFAENSVPVINAGDAGOHPTOTLIDL-----YT 142
 QY 196 FSNFIGLESRIISQGAATTSISDAK--RAVDFILLYCOLVYMR-----ETLLV 241
 DB 143 IKKEGRLD-----GITIALMGDLKYSRTIHSILKALALPDMKILYLSPEALALPEIIE 197
 QY 242 DLAILYKRGNAEHVASAVE--NANRYNKELAADTLDFLHKL-----IPEQALIGA----- 289
 DB 198 DVSAEIRARLBEVIEIDLVYTRIOKEEFPEDEEY-RKVSQSYRTAETLKSADSMI 256
 QY 290 VYHPI-----SASESKALINMYKRGVP 313
 DB 257 VMHPLPRVDEIHPSVDSTKHARYFOQSFTGVP 288

RESULT 12
 HTR1_HAINT STANDARD; PRT; 535 AA.
 ID AC P33741; 09HPF6;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE SENSOR RHO DOPSIN I TRANSDUCER (HTR-1) (METHYL-ACCEPTING PHOTOTAXIS PROTEIN I) (MPP-I).
 GN HTR1 OR HTR1 OR VNG1659G.
 OS Halobacterium sp. (strain NRC-1), and
 OS Halobacterium halobium.


```

OC Archaeae;Euryarchaeota; Halobacteriales; Halobacteriaceae;
CC Halobacterium.
OX NCBI_TaxID=64091, 2242;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NR-1;
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairs G.G., Bergquist B., Pan M.,
RA Shukla H.D., Laszky S.R., Balliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welter R., Goo Y.A.,
RA Lelthausen B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angvine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudis J.L., Jung K.-H.,
RA Alen M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RA "Genome sequence of Halobacterium species NRC-1."
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-10; 350-377 AND 457-476.
RC SPECIES=H. halobium; STRAIN=FLXSR;
RX MEDLINE=93101637; PubMed=1465418;
RA Yao V.J., Spudis J.L.;
RT "Primary structure of an archaeobacterial transducer, a
RL methyl-accepting protein associated with sensory rhodopsin I."
RL Proc. Natl. Acad. Sci. U.S.A. 89:11915-11919(1992).
CC -1- FUNCTION: TRANSDUCES SIGNALS FROM THE PHOTOTAXIS RECEPTOR
CC SENSORY RHODOPSIN I (SR-I) TO THE FLAGELLAR MOTOR. RESPONDS TO
CC LIGHT CHANGES THROUGH THE VARIATION OF THE LEVEL OF METHYLATION.
CC -1- SIMILARITY: BELONGS TO THE CHEMOTAXIS SENSORY TRANSDUCERS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as their content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licenselsb-sib.ch).
CC -----
DR EMBL, AE005075; AAC19913.1; -
DR EMBL, L05603; AAA72315.1; -.
DR PIR, A47190; A47190.
DR InterPro: IPR0001022; Chemotaxis_transducer.
DR InterPro: IPR000658; DUF5
DR InterPro: IPR003660; HAMP.
DR Pfam: PF00672; DUF5; 1.
DR Pfam: PF00015; MCPSignal; 1.
DR SMART: SM00304; HAMP; 2.
DR SMART: SM00283; MAP; 1.
KW Transducer; Photoreceptor; Transmembrane; Methylation.
FT INTL_MET 0
FT DONAIN 1 13 CYTOPLASMIC (POTENTIAL) .
FT TRANSMEM 14 28 POTENTIAL.
FT DONAIN 29 38 EXTRACELLULAR (POTENTIAL) .
FT TRANSMEM 39 54 POTENTIAL.
FT DONAIN 55 535 CYTOPLASMIC (POTENTIAL) .
FT MOD_RES 265 265 METHYLATION.
FT MOD_RES 272 272 METHYLATION.
FT MOD_RES 279 279 METHYLATION.
FT MOD_RES 463 463 METHYLATION.
FT MOD_RES 472 472 METHYLATION.
FT SEQUENCE 535 AA; 56544 MW; B9945E4F66A9D091 CRC64;

```

Query Match	4.3%	Score 101	DB 1	Length 555
Best Local Similarity	20.3%	Pred. No 9.9%		
Matches	59	Conservative	57	Mismatches 117; Indels 58; Gaps 11;
OY	25	DVNSLITVEYFALKASGSENEALEGLK-----GEIQTKPDYRGQATKILGSGS	77	
		: : : : : : : : : : : : : : : : : : : : : : : : :		
Db	272	DLASIEDVYATAASPDARSSKSALEDEMSIETFEVDVAGVYEQULDQVAETRIDYDVI--	329	
		: : : : : : : : : : : : : : : : : : : : : : : : :		
OY	78	ALGRKNSGDARKIISGCDIYAGITFTTGPGPGMGKIGAVASVSSILSLFTGSSAKNSYA	137	
		: : : : : : : : : : : : : : : : : : : : : : : : :		

```

Db      330  ----TDIGEOINMLA--LMSIEPAARGNADGSEVADEVKDLAE--ETODRANEIA 361
OY      138  AVIDRALSKRHD--EAIORHAAGAKRPEAESAFIOVMKOOSNLTDSDLSIIANMPYVK 195
        : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      382  AVEVEKTAQTEDYASTIO-----TRRVEGCS-----ETVESTLMDINTIADSLA--E 428
OY      136  FSNFIGOLESLISOGAAT-----TSLSAKRAVDFILLYCOLVVMRETLVYDLA 244
        : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      429  VSNSTDELQRTTSGOAEFVSTATSEVERAGLSDDTTAL-----ASDAESA 474
OY      245  ILYRKGNAEHVASAVEN--ANFVKNELADTLDLPHKLIPCALIGAYVHPI 294
        : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      475  VIGRRESAEETIASLEQFOPTAVYQLOSRVASFVTAETDESETAGSGVEQPV 525

```

```

RESULT 13
YABE_BACSU          STANDARD;          PRT;          437 AA.
AC      P37546;
DT      01-OCT-1994 (Rel. 30, Created)
DT      01-OCT-1994 (Rel. 30, Last sequence update)
DT      20-AUG-2001 (Rel. 40, Last annotation update)
DE      HYPOTHEICAL 47.7 KDA PROTEIN IN METS-MSGA INTERGENIC REGION.
YABE.
OS      Bacillus subtilis.
OC      Bacteria; Firmicutes; Bacillus/Clostridium group;
OC      Bacillus/Staphylococcus group; Bacillus.
OX      NCBI_TaxID=1423;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=168;
RX      MEDLINE=96051385; PubMed=7584024;
RT      Ogasawara N., Nakai S., Yoshikawa H.;
RL      "Systematic sequencing of the 180 kilobase region of the Bacillus
      subtilis chromosome containing the replication origin.";
      DNA Res. 1:1-14(1994).
-----
CC      This SMIS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
CC      EMBL; D26185; BAA05275.1; -.
DR      EMBL; Z99104; CAB1816.1; -.
DR      Subtilist; BG10103; yabE.
KW      Hypothetical protein; Complete proteome.
SQ      SEQUENCE 437 AA; 47710 MW; E1B957CDED7FF94A CRC64;

```

[illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 12:11:47 ; Search time 19.46 Seconds

(without alignments)
1761.488 Million cell updates/sec

Title: US-09-647-522-5

Perfect score: 2322
Sequence: 1 MILKHLPLFLFVLAITSAKH.....SGYRSKNNPPGQGHMST 450Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2317	99.8	450	2 JC7371	toxin - jellyfish
2	992.5	42.7	463	2 JC7372	toxin-A - jellyfish
3	137	5.9	1222	2 G59100	hypothetical prote
4	117	5.0	833	2 S54776	transcriptional fac
5	116.5	5.0	423	1 DEBYMC	malate dehydrogena
6	116	5.0	530	2 C72291	methy1-accepting c
7	112.5	4.8	510	2 C84718	probable kinesin 1
8	112.5	4.8	964	2 E71460	probable outer mem
9	111.5	4.8	656	2 A72428	methy1-accepting c
10	111.5	4.8	1336	2 T18288	ABC transport prot
11	110.5	4.8	656	2 E72379	methy1-accepting c
12	110.5	4.8	661	2 G72316	methy1-accepting c
13	110.5	4.8	1024	2 S10056	hemolysin A - Esch
14	110	4.7	541	2 B84185	cytochrome-like pr
15	109	4.7	541	2 B83032	probable chemotaxi
16	109	4.7	612	2 B75282	probable acyl-CoA
17	108.5	4.7	659	2 G82365	methy1-accepting c
18	108.5	4.7	1023	1 LEECA	hemolysin A - Esch
19	106.5	4.6	1098	2 B70232	hypothetical prote
20	106	4.6	1975	2 B81192	hemagglutinin/hemo
21	105.5	4.5	312	2 C55578	hypothetical prote
22	105	4.5	506	2 E65095	aerotaxis receptor
23	105	4.5	506	2 A85968	aerotaxis sensor r
24	103	4.4	1456	2 S14005	hypothetical prote
25	102.5	4.4	460	2 T19111	hypothetical prote
26	102	4.4	1456	2 A37853	flagellin, 40k - p
27	102	4.4	1456	2 JQ2294	hypothetical 16S.1
28	101.5	4.4	299	2 B69263	aspartate carbamoy
29	101.5	4.4	428	2 D81255	histidinol dehydro

30	101	4.3	490	2 T00404	probable cytochrom
31	101	4.3	536	1 A47190	transducer protein
32	101	4.3	536	2 E84318	htr1 transducer li
33	101	4.3	1520	2 B82274	conserved hypotnet
34	100.5	4.3	307	1 DEPGC	3-hydroxyacyl-CoA
35	100.5	4.3	548	2 S15910	hemolysin-releasin
36	100.5	4.3	909	2 A75337	exonuclease Shbc -
37	100	4.3	437	2 S66069	conserved hypotnet
38	100	4.3	444	2 E72031	conserved hypotnet
39	100	4.3	444	2 F86591	C7579 hypothetical
40	99.5	4.3	1510	2 T13634	probable minor tai
41	99	4.3	484	2 E84254	hypothetical prote
42	99	4.3	931	2 F84637	probable plasma me
43	98.5	4.2	379	2 G82107	flagellin core pro
44	98.5	4.2	795	2 S26712	hypothetical prote
45	98.5	4.2	1517	2 T13329	hypothetical prote

ALIGNMENTS

RESULT	1	Score	DB 2	Length
JC7371	toxin - jellyfish (Carybdea rastoni)	99.88	2317	450
C:Species: Carybdea rastoni				
C:Date: 17-Nov-2000				
C:Accession: JC7371				
C:Accession: JC7371				
R:Nagai, H.; Takawa, K.; Nakaio, M.; Ito, E.; Miyake, M.; Noda, M.; Nakajima, T.				
Biochem. Biophys. Res. Commun. 275, 582-588, 2000				
A:Title: Novel proteinaceous toxins from the box jellyfish (sea wasp) Carybdea rasto				
A:Reference number: JC7371				
A:Accession: JC7371				
A:Molecule type: mRNA				
A:Residues: 1-450 <NAG>				
A:Cross-references: DDBJ:AB015878				
A:Accession: PC7094				
A:Molecule type: Protein				
A:Residues: 39-55;56-70;196-210;250-267;268-279;309-325;363-377;378-382 <NA>				
C:Comment: This protein, a member of bioactive protein, has hemolytic activity.				
C:Keywords: hemolysis; inflammation; toxin				
Query Match	99.88	Score 2317	DB 2:	Length 450
Best Local Similarity	99.88	Pred. No. 8.9e-155		
Matches 449; Conservative 0; Mismatches 1; Indels 0; Gaps 0;				
QY 1 MILKHLPLFLFVLAITSAKHGRSDVNSLTRKVFETALKEASGSNEALEGLKEGTOT 60				
Db 1 MILKHLPLFLFVLAITSAKHGRSDVNSLTRKVFETALKEASGSNEALEGLKEGTOT 60				
QY 61 KPDVVGATKILGSGVSGALGKLSGDAATKIIISGLDIYAGIATTFGGVGMGIGAVASPV 120				
Db 61 KPDVVGATKILGSGVSGALGKLSGDAATKIIISGLDIYAGIATTFGGVGMGIGAVASPV 120				
QY 121 SSIIISLFGSSAKNSVAVIDRALSKHDEAIORHAAGAKRDFESSAFIOVMKOOSNLT 180				
Db 121 SSIIISLFGSSAKNSVAVIDRALSKHDEAIORHAAGAKRDFESSAFIOVMKOOSNLT 180				
QY 181 DSDSIITANPVYKFSNFISQLESRIISOGAATTSISDAKRAVDFITLLYCOLVVRRETL 240				
Db 181 DSDSIITANPVYKFSNFISQLESRIISOGAATTSISDAKRAVDFITLLYCOLVVRRETL 240				
QY 241 VDLAILYRKGAHVAHAVENANRVNKLADTLDFLKLIPEDALIGAVYHPISASSETS 300				
Db 241 VDLAILYRKGAHVAHAVENANRVNKLADTLDFLKLIPEDALIGAVYHPISASSETS 300				
QY 301 KALINMYRKEGVPPVPRIGNRKRYKFTNSYNTYSISCEAYMGVYMPFGCSNVNPNRIY 360				
Db 301 KALINMYRKEGVPPVPRIGNRKRYKFTNSYNTYSISCEAYMGVYMPFGCSNVNPNRIY 360				
QY 361 SKMSDGEYTMENSDRRKLYIRKHQDGWGWGLDDEPDGQGHMREIPLRHGYVYSSKRW 420				
Db 361 SKMSDGEYTMENSDRRKLYIRKHQDGWGWGLDDEPDGQGHMREIPLRHGYVYSSKRW 420				

OY 421 NMFYMESSASGYIRSWENNPQGHWSIT 450
 DB 421 NMFYMESSASGYIRSWENNPQGHWSIT 450
 RESULT 2
 JC7372
 toxin-A - jellyfish (Carybdea alata)
 C.Species: Carybdea alata
 C.Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 17-Nov-2000
 C.Accession: JC7372; PC7095
 R.Nagai, H.; Takuwa, K.; Nakao, M.; Sakamoto, B.; Crow, G.L.; Nakajima, T.
 Biochem. Biophys. Res. Commun. 275, 589-594, 2000
 A.Title: Isolation and characterization of a novel protein toxin from the Hawaiian box
 A.Reference number: JC7372
 A.Accession: JC7372
 A.Molecule type: mRNA
 A.Residues: 1-463 <NA>
 A.Cross-references: DDBJ:AB036714
 A.Accession: PC7095
 A.Molecule type: protein
 A.Residues: 158-168;431-463 <NA>
 A.Comment: this protein, belongs to a class of bioactive proteins with hemolytic activity
 C.Keywords: hemolysis; inflammation; toxin

Query Match 42.7%; Score 992.5; DB 2; Length 463;
 Best Local Similarity 43.5%; Pred. No. 5,4e-62;

Matches 197; Conservative 95; Mismatches 148; Indels 13; Gaps 7;

OY 9 LFLVLAITSKAGH-----KRSVDNSLITFVETALKEAS--GSNEAL-FALEGLKEIOT 60
 DB 11 LFLVSTAPPSQARLSRYRSADAVSTDIDGILGDLNDLGTTRKLKELGCVQAVK 70
 OY 61 KP-DRYGQATKILGVSALGKINSQDAFKIISGLDIAGIATTEGPGVGMIGAVASF 119
 DB 71 EPATTAIKVSTIVSGVSLKFKSGDPDPAVSGCLDIASVATPGEGYIGAVASV 130
 OY 120 VRSISLTFGSSAKNSVAIVFRLSKHDEATORHAGAKRDPFASSAFIOVMKOOSNL 179
 DB 131 ISSIISLFGNSMSAIKOVIDDAFKKRDQEDNVKQAKKRFNAVITFVNSSTENL 190
 OY 180 TSDLSIIANYPVYKFSNFIGLESRIISQGAATFSLDAKRAVDLILYCOLVYARETL 239
 DB 191 TEVHLDSDVAVDAVFATMLGVLSESRINRGYSVTDNEMARTINFIPLYDLSVARETL 250
 OY 240 LYDLALILYK--GNAEHVASAVENANRVKELADTLDELHLKILPEQALIGAVYHPISAS 297
 DB 251 LVQVILLYRAGAVDELALSLTSDONKEATRETVEFLHOMETKYSICGSIYYPIIDS 310
 OY 298 ETSKALINTKTFGVDPVR-PIGNRRYKFTSNYNTYISCSFAYVGNMFRGCSNVRNP 356
 DB 311 KALIGILKLTKEFGVDPARYTFDGLYRMONRAMRYSICESTAGNMFRCQKDSSTH 370
 OY 357 NRVSKMSDGYFTMNSDRKRLYTKHDSQMGWGLDEDDPGDGHMRTPLRHKYVSS 416
 DB 371 GRIKIKLENGYHTI-PLRSKAMVYTHAGQWGWGTDEDDPGEGYTFITPLNGFPMVST 429
 OY 417 KRPWPMFMMESSASGYIRSWENNPQGHWSI 449
 DB 430 KRPMPDFYVMESSAHGYIRSMYHNPDPQOMKI 462

RESULT 3
 G59100

hypothetical protein PX01-79 - Bacillus anthracis virulence plasmid PX01
 C.Species: Bacillus anthracis

C.Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 16-Feb-2001

C.Accession: G59100

J.Okunaka, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hill, K.K.; Keim, P.; Koehler
 J. Bacteriol. 181, 6509-6515, 1999
 A>Title: Sequence and organization of PX01, the large Bacillus anthracis plasmid harbor

A.Reference number: A59091; MUID:99445483
 A.Accession: G59100
 A.Status: preliminary
 A.Molecule type: DNA
 A.Residues: 1-1222 <OK>
 A.Cross-references: GB:AF065404; NID:g4894216; PIDN:AAD2383.1; PID:g4894295
 A.Experimental source: strain Sterne
 A.Note: similar to hypothetical, hydrophobic protein (567 aa), Bacillus firmus (U645;
 C.Genetics:
 A:gene: PX01-79
 A:genome: plasmid
 C:superfamily: hypothetical protein PX01-79

Query Match 5.9%; Score 137; DB 2; Length 1222;

Best Local Similarity 23.0%; Pred. No. 0.17;

Matches 107; Conservative 55; Mismatches 173; Indels 130; Gaps 23;

OY 21 GKRSDV-NSLITFVETALKEASGSNEALFALEGLKEIOTKPDVGOATKIL--GSVGS 77
 DB 310 GVRDDYRNSKVRKPDGDKDEK--TKEDADNPKALANLGTDIGTTSRAARMKAGOVGS 367
 OY 78 ALGKNSGDATKIISGLDIAVGIATTFGPGVGMIGAVASFVSSILSLFTGSSAKNSVA 137
 DB 368 ALGKATTT-----LAGMAA--GAGLGPAMVAGSGVSGALGAAPGAAGRSVA 413
 OY 138 AYIDR--ALSKHDEAIORHAGAKRDPFASSAFIOVMKOOSNLTDSDLS--IIANVPV 193
 DB 414 AVGEAGVALAGKHGSSIKKGADYSNLNGES-----YPELTDEDIAODLATPDFEN 464
 OY 194 YFNSNFIGLESRIISQGAATFSLSD-----AKRAVDLILYLC-----OLVYARETL 243
 DB 465 WKADNPNSAVASRLKQAFNPASDAELAKVAKTNSQMSRFTQKRODLONNKKT----- 519
 OY 244 AILYRKGNAEHVASAVENANRVKELADTLDELHLKILPEQALIGAVYHPISASETSK-- 301
 DB 520 ATPY--GNARDLVNATNA--FOKGYEGDHKPTFMSQLPEN-----MSAEKEKQM 566
 OY 302 -----AILNY-----TKYFGVP-DVPRPIGNRRYKFTSNYNTYISCSFAYVGNM 346
 DB 567 NDLNAKVOGFRNHAEQATKAGAMPVDAKDKOGNN--LFDKSYVKKDPAFASOLAG--- 621
 OY 347 FRGCSNVRNPNIKVSMSDGF--YTMENSDR-----RLYTKHDSQMG 388
 DB 622 -----KVGAVAGVAGGETLESQHONNGVAGALVGASTAFKKGYTDHKAQF-- 668
 OY 389 WGTLEDDEPDGQHMFILPLRHGKYVSSKRPWPMFMMESSASGY 433
 DB 669 ---MKQFPAD-----MSQDEKEAMNKHLDQKVOGF 696

RESULT 4

S54776
 transcription factor - mouse

C.Species: Mus musculus (house mouse)

C.Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 05-Nov-1999

C.Accession: S54776

R.Evers, R.; Smid, A.; Rudloff, U.; Lottspeich, F.; Grunmt, I.

EMBO J. 14, 1248-1256, 1995
 A>Title: Different domains of the murine RNA polymerase I-specific termination factor

A.Reference number: S54776; MUID:95337206

A.Accession: S54776

A>Status: preliminary

A.Molecule type: mRNA

A.Residues: 1-833 <EVE>

A.Cross-references: EMBL:X83974; NID:g639818; PIDN:CAA58808.1; PID:g639819

Query Match 5.0%; Score 117; DB 2; Length 833;

Best Local Similarity 22.0%; Pred. No. 2.5;

Matches 100; Conservative 54; Mismatches 127; Indels 174; Gaps 23;

OY 4 KHL----PMLFVLAITSKAGKRSVDNSLITFVETALKEASG-SNEALFALEGLK--- 55

```

Db      235 KHLQKVAW-----DVVQSGSPESISLPSEPLSSEDLCKSTEAATFCRSLKKNV 286
               |||
               |||
               |||
OY      56 ---GEIQTKPRYGAQTKILGVSALGKNSDARKITISGCDIYAGIATTTGGPYGMG 112
               |||
               |||
               |||
Db      287 FRSGELEIPD-----SLDDETTISERLD-----STHHGAVGAG 321
               |||
               |||
               |||
OY      113 -----IGAVASFVSSILSLFTGSSAKNSVAVIDR-----142
               |||
               |||
               |||
Db      322 ECSTKESHISIKKKSKKKHKSVALATSSDSASVYTSKAKN---ALVDSSESGAVREED 378
               |||
               |||
               |||
OY      143 -----ALSKHDEAIOI---HAACAKRDEAFESSAFIQMKQOQSNLTDSDLS 185
               |||
               |||
               |||
Db      379 VDHREAEAAQACSTEKHR-EMQRFLEPHEESNESASNSAARISDSRSDSDVD 437
               |||
               |||
               |||
OY      186 IIAANPVKFSNFIGULGSLRSQGAATSLSDAKRAVFIILYCOLVYVRETLVLDLAI 245
               |||
               |||
               |||
Db      438 LQSA---VAQREFIPDIOER---AATV-----IRRWYRDDLG- 469
               |||
               |||
               |||
OY      246 LYRKGNAEHVA-----SAVENANRVNKEADTLDFLKLIPEDQALIGAVVHPSASST 299
               |||
               |||
               |||
Db      470 LFEKKAQCVAIRFGKFSKE-----NKOIEKNVODFL-----SLTGIES 509
               |||
               |||
               |||
OY      300 SKAIIATYKTFEGVDPVPRPIGN--RRYKFTSYNWTYSICSEAYMGVNFRCGSMVRNPN 357
               |||
               |||
               |||
Db      510 ADKLITYTRY---PEEKLITLTKRKHAFL-----RLHIGKGIARWPKLY---Y 551
               |||
               |||
               |||
OY      358 IRVSKMSD-----GFYTMENSDRKLYITTKHDQGM 387
               |||
               |||
               |||
Db      552 YRAKIFDVNNYKGRYNEEDFKKLAYHSLHGNDW 586
               |||
               |||
               |||

```

RESULT 5
DEBYMC
malate dehydrogenase (EC 1.1.1.37), cytosolic - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: protein AOE23; protein 00537; protein YOL126C
C:Species: *Saccharomyces cerevisiae*
C>Date: 30-Jun-1993 #sequence,revision 19-Jul-1996 #text,change 16-Jun-2000
C/Accession: S63444; S12937; A34986; S05770; S66823; S71982
R:Casamayor, A.; Khalid, H.; Balcells, L.; Aldea, M.; Casas, C.; Herrero, E.; Arino, J.
Submitted to the EMBL Data Library, November 1995
A:Description: Sequencing of a 13.4 kbp fragment of the left arm of chromosome XV reveal
A:Reference number: S63440
A:Accession: S63444
A:Molecule type: DNA
A:Residues: 1-423 <CNS>
A:Cross-references: EMBL:U41293; NID:g1209710; PIDN:AAC49466.1; PID:g1209715
A:Experimental source: strain FY1679
R:Minard, K.I.; McAlister-Henn, L.
Mol. Cell. Biol. 11, 370-380, 1991
A:Title: Isolation, nucleotide sequence analysis, and disruption of the MDH2 gene from *S.
A:Reference number: S12937; MUID:91094852*
A:Accession: S12937
A:Molecule type: DNA
A:Residues: 47-423 <MIN1>
A:Cross-references: EMBL:M62808; NID:g171915; PIDN:AAA34766.1; PID:g171916
A:Accession: A34986
A:Molecule type: protein
A:Residues: 48-73 <MIN2>
R:Kopetzki, E.; Entian, K.D.; Lotspeich, F.; Mecke, D.
Biochim. Biophys. Acta 912, 398-403, 1987
A:Title: Purification procedure and N-terminal amino acid sequence of yeast malate dehy
A:Reference number: S05770; MUID:87185517
A:Accession: S05770
A:Molecule type: protein
A:Residues: 48-73, P, 75-81 <KOP>
R:Arino, J.; Casamayor, A.; Gamio, F.J.; Gancedo, C.; Iafuente, M.J.; Aldea, M.; Casas, C
submitted to the Protein Sequence Database, July 1996
A:Reference number: S66823
A:Accession: S66823
A:Molecule type: DNA
A:Residues: 1-423 <ART>
A:Cross-references: EMBL:274868; NID:g1420008; PIDN:CAA99145.1; PID:g1420009; GSPDB:GN00

```

A:Experimental source: strain S288C  
R:Casamayor, A.; Khalid, H.; Balcells, L.; Aldea, M.; Casas, C.; Herrero, E.; Arino  
Yeast 12, 1013-1020, 1996  
A:Title: Sequence analysis of a 13.4 kbp fragment from the left arm of chromosome X  
ending frames.  
A:Reference number: S71978; MUID:97051588  
A:Accession: S71982  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-423 <CANP>  
A:Cross-references: EMBL:U41293; NID:g1209710; PIDN:AAC49466.1; PID:g1209715  
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, November 19  
C:Genetics:  
A:Gene: SGD:MDH2; MIPS:YOL126C  
A:Cross-references: SGD:S0005486; MIPS:YOL126C  
A:Map position: 15L  
C:Superfamily: L-lactate dehydrogenase  
C:Keywords: cytosol; homodimer; NAD; oxidoreductase  
F:48-423/Product: malate dehydrogenase #status experimental <MAT>  
F:228,261/Active site: Asp, His #status predicted  
F:231/Binding site: substrate (Arg) #status predicted

```

Query Match 5.0%; Score 116.5; DB 1; Length 423;
Best Local Similarity 20.2%; Pred. No. 1.1;
Matches 72; Conservative 65; Mismatches 125; Indels 95; Gaps 17;

```

OY      13 LAITSAKHGKSDVNSL-TVEYALKEASGS-----NEALEALEGKGEIOTRPD 63
               |||
               |||
               |||
Db      62 IAILGAAGIGQSSLSLTKAQLQYOLKESNRSVTHILALVDVQNEALINGVATDLSHID 121
               |||
               |||
               |||
OY      64 RVQCATKILGVSALGKLGNGDAPTKIISG-----CLDIAGIATTTGGPYGMG 112
               |||
               |||
               |||
Db      122 PIVSSH--SPAGCIENCLHNASTIVIPAGVPRPGMTRODLFVFNWAGIISQDLSI--- 176
               |||
               |||
               |||
OY      113 IGAVASFVSSILSLFTGSSAKNSVAVIDRALSKH---RDEAIOIRHAG-AKRDFAESSA 168
               |||
               |||
               |||
Db      177 --AECCLSKFEVLVINPNVSLVPVAVNSLTKHHPNSNGIRRMIGVTKLDIVRAST 234
               |||
               |||
               |||
OY      169 FIVMKQOQSNLTD-----SDLSITAN-----VPVYKFSNFIGQL----- 203
               |||
               |||
               |||
Db      235 FLKEINIESGLTPRVNSMPVPVIGSHGCEITIPLFQSNFLSLNBDQLKYLHROYG 294
               |||
               |||
               |||
OY      204 -----ESRISQGAATTSLSDA--KRAVDFTLLKQLYVMETLLVDLAILRKGNAEHVA 256
               |||
               |||
               |||
Db      295 GDEVYAKNKGKSNATISMAGTKCVQEV-----SLL-----GNIEQIH 335
               |||
               |||
               |||
OY      257 SA---VENANRVNKEADTLDFLKLIPEDQALIGAVVHPSASSTSKAIIANTYKY 309
               |||
               |||
               |||
Db      336 GTYYVPLKDNANPPIAPGAD-----QLDP--LVGDGADYFAIPLTITTKGV-STVDY 383
               |||
               |||
               |||

```

RESULT 6
C72291
methyl-accepting chemotaxis protein - *Thermotoga maritima* (strain MSB8)
C:Species: *Thermotoga maritima*
C>Date: 11-Jun-1999 #sequence,revision 11-Jun-1999 #text,change 21-Jul-2000
C/Accession: C72291
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hic
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,
C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
A:Reference number: A72200; MUID:99287316
A:Accession: C72291
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-530 <ARN>
A:Cross-references: GB:AF001771; GB:AE000512; NID:g4981678; PIDN:AAD36219.1; PID:g498
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM1143
C:Superfamily: methyl-accepting chemotaxis protein

Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.
C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316
A:Accession: A72428
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-656 <ARN>
A:Cross-references: GB:AE001690; GB:AE000512; NID:94980496; PIDN:AAD35117.1; PID:9498050
C:Genetics:
A:Gene: TM0023
C:Superfamily: probable methyl-accepting chemotaxis transducer

Query Match
Best Local Similarity 4.8%; Score 111.5; DB 2; Length 656;
Matches 63; Conservative 44; Mismatches 107; Indels 95; Gaps 10;

QY 24 SDVNSLTKEVETALKESGSENALEAEGT--KGEIOTKPDVGOATKIIGSVGSALGK 81
Db 393 NMSSALTEVTSVGEVVAASQNIKITODTERSEAVTKAARG--TERVEAVGVINK 450
QY 82 LNSG-----DAKTIISGCLDIYAGIA-----TTFGPGVGMIGAV 116
Db 451 LKSAERORDYLRLEVDASAKTIGEIVDTISSIAEQTNLALNAIEAARGEGRGFAVY 510
QY 117 ASPVSSILSTFGSSAKNSVAVIDRALSKHDEAIOHNAAGKDFPAESSAFIOWKQO 176
Db 511 ADEIRKL-----AEESORATEDIAKMLSSIRATIEHVENSGKEMFGVDEIANMGEE 562
QY 177 SNLTDSLSITIAANVPYKFSNFIOLE---SRISOGAATTSLSDAKRAVDFTILYCOLV 233
Db 563 -----VTKRFREILGRIEEINSMIENTAATAOEGAA----- 594
QY 234 VMRETLVDLAILYRKGNAEHVAASAVENANVKNELAADTDLFKLIPEDALLGAVYHP 293
Db 595 -----AEEMASAMDNVTYKI-VEGVESLNRMESLIEDOT----- 627
QY 294 ISASETSKA 302
Db 628 ESAARVSEA 636

RESULT 10
T18288
ABC transport protein - slime mold (Dictyostelium discoideum) (fragment)
C:Species: Dictyostelium discoideum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
R:Loomis, W.F.
submitted to the EMBL Data Library, August 1996
A:Reference number: Z18857
A:Accession: T18288
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1336 <LOO>
A:Cross-references: EMBL:066526; NID:91513297; PID:91513298; PIDN:AAB06789.1
C:Genetics:
A:Gene: abca

Query Match
Best Local Similarity 4.8%; Score 111.5; DB 2; Length 1336;
Matches 66; Conservative 49; Mismatches 112; Indels 91; Gaps 14;
QY 7 PWLFIV-----LAITSAKI-----GKRSVNSLTKEVETALKESGSENALEAEGT 56
Db 334 PWFFVLPYWGSLTKVKKHIPPYFDDEVDVRAAITKADHA-----SNRAPL-IICGLSK 386
QY 57 EIQT--KPDVGOATKIL-----GSVGSALGKLNSSGATKI--ISGCLDIYAGIATTFG 106

Db 387 SYTKLFRKKTVHAKVYLSLSEVEKGTILIGLGSNGCGKSTTIGMLTGLLEPTAGDALYGG 446
QY 107 GPVGMIGAVASFEVSSI-----LSLEF-----GSSAKNSVAAYIDRA-L 144
Db 447 HSYISNTAAVRRITTSVPPQHDILMAEMTAREHQLFSELGIPAOERESQIOKVLQVRL 506
QY 145 SKHREDAIORHAAKAR-----DFAESSAFIOWKQO-- 178
Db 507 SKISNNLSTISGCKMRRLSVALACIGDPKIFPMDERTTGVDSKRHLIDLKYSIKNDK 566
QY 179 ---LSDSDI-----STIAANVPYKFSNFIOLESRSISOGAATTSLSDAKRAVDFTIL 227
Db 567 VILTSHDMHEVELADKIVINMEGVAMCNGNSLOLKSXYGQSYNIYAKSPESIPAVY 626
QY 228 LYCOLV-----MRETL 240
Db 627 EFVTLSPCKFMKOSAL 644

RESULT 11
E72379
methyl-accepting chemotaxis protein - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: E72379
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.;
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardsc
C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from geno
A:Reference number: A72200; MUID:99287316
A:Accession: E72379
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-656 <ARN>
A:Cross-references: GB:AE001721; GB:AE000512; NID:94980922; PIDN:AAD35514.1; PID:94
C:Genetics:
A:Gene: TM0429
C:Superfamily: probable methyl-accepting chemotaxis transducer

Query Match
Best Local Similarity 4.8%; Score 110.5; DB 2; Length 656;
Matches 62; Conservative 45; Mismatches 107; Indels 95; Gaps 10;

QY 24 SDVNSLTKEVETALKESGSENALEAEGT--KGEIOTKPDVGOATKIIGSVGSALGK 81
Db 393 NMSSALTEVTSVGEVVAASQNIKITODTERSEAVTKAARG--TERVEAVGVINK 450
QY 82 LNSG-----DAKTIISGCLDIYAGIA-----TTFGPGVGMIGAV 116
Db 451 LKSAERORDYLRLEVDASAKTIGEIVDTISSIAEQTNLALNAIEAARGEGRGFAVY 510
QY 117 ASPVSSILSTFGSSAKNSVAVIDRALSKHDEAIOHNAAGKDFPAESSAFIOWKQO 176
Db 511 ADEIRKL-----AEESORATEDIAKMLSSIRATIEHVENSGKEMFGVDEIANMGEE 562
QY 177 SNLTDSLSITIAANVPYKFSNFIOLE---SRISOGAATTSLSDAKRAVDFTILYCOLV 233
Db 563 -----VTKRFREILGRIEEINSMIENTAATAOEGAA----- 594
QY 234 VMRETLVDLAILYRKGNAEHVAASAVENANVKNELAADTDLFKLIPEDALLGAVYHP 293
Db 595 -----AEEMASAMDNVTYKI-VEGVESLNRMESLIEDOT----- 627
QY 294 ISASETSKA 302
Db 628 TSAARVSEA 636

RESULT 12
G72316

[illegible]

Db 342 DHDAVEEGSEVLAQENITRAGSGVSDALSOVDNVLVAGIKALTTDFEAVILETGRAL 401

Oy 240 LVDLAILYRKGNAPHVASVE 260

Dd 402 LGDAREQYHRCNAERAAASVE 422

RESULT 15

Probable chemotaxis transducer PA4915 [imported] - *Pseudomonas aeruginosa* (strain PA01)
 C:Species: *Pseudomonas aeruginosa*
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: E83002
 R:Stovey, C.K.; Plam, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; B
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim,
 S.; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
 A:Reference number: A82950; MUID:20437337
 A:Accession: E83032
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-541 <SMO>
 A:Cross-references: GB:AE004904; GB:AE004091; NID:g9951184; PIDN:AAC08300.1; GSPDB:GN001
 A:Experimental source: strain PA01
 C:Genetics:
 C:Gene: PA4915

Query Match	4.78;	Score 109;	DB 2;	Length 541;
Best Local Similarity	22.6%;	Pred NO 5 1;		

Best Local Similarity 22.6%; Pred. No. 5.1;
Matches 64; Conservative 45; Mismatches 94; Indels 80; Gaps 13;

QY	2	ILKHFWLFIYATLS-----AKHKRSD-VNSLITKYETALTAKKASGNEA	46
Db	262	LIRHIAETARKLITISDEKSAISNOVSHVATSGDDTDMDATAVHOSMAVODVARNQA	321
QY	47	ALEALEGLKEIQT-----KPDVGGATKILISVSGALCKLNSGATKILISGCLIV-	98
Db	322	AAASAESANSAHGTGLVHANLDAIQLSYVWGEAGAVIDTLR--NKTDEISIVLEVIQ	379
QY	99	-----AGATPFGGPVMGIGAVASFSS-----ILSLPTGSS	131
Db	380	NIAGOTNLALMAAIEARAGEAGRGRAVVADEVRSLATYTHKATETIREITEALQAGAS	439
QY	132	AKNSVAAYIDRALSCKHDEA---IQR-HAAG-----AKRDFAESAFIYQVAKQSNL	179
Db	440	SAVSF-----MOOSEQAOVSVORAHNEGAKGLAQAVEGIAQSNQASISATEEOTA	492
QY	180	TDSDLSIIAANVPYKRSNFGOLESLISOGAATSLISDAPKA	222
Db	493	TASEVS---QNIDSLANS--IGF---VACAVKTSITSVELA	526

Search completed: May 8, 2002, 12:13:54
Job time: 127 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 12:11:27 ; Search time 15.85 seconds
(without alignments)
638.895 Million cell updates/sec

Title: US-09-647-522-5

Perfect score: 2322

Sequence: 1 MILKHPMLFIVLATSARKH.....SGYIRSWENNPQGHWSIT 450

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

212252

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/prodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/prodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/prodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/prodata/2/1aa/6C.COMB.pep.*
6: /cgn2_6/prodata/2/1aa/6D.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	4.7	3031	1	US-07-689-008-2
2	100.5	4.3	3079	4	US-09-413-814-80
3	97	4.2	2123	4	US-08-968-685A-10
4	95	4.1	434	2	US-08-989-925-1
5	95	4.1	1456	1	US-08-803-973-2
6	95	4.1	1456	1	US-08-803-972-2
7	94.5	4.1	1489	6	5183745-2
8	94.5	4.1	1794	6	5183745-6
9	93.5	4.0	580	2	US-08-591-079-2
10	93	4.0	1296	1	US-08-480-604A-28
11	93	4.0	1296	1	US-08-405-496A-28
12	93	4.0	1296	4	US-08-915-136-28
13	92.5	4.0	432	2	US-08-705-660-18
14	92.5	4.0	432	3	US-08-989-045-18
15	92.5	4.0	699	3	US-08-694-865-16
16	92.5	4.0	699	3	US-09-124-491-16
17	92	4.0	497	4	US-08-740-223A-14
18	92	4.0	593	2	US-08-591-079-8
19	92	4.0	593	2	US-08-591-079-10
20	91.5	3.9	491	4	US-08-740-223A-13
21	91.5	3.9	937	1	US-08-253-155A-31
22	91	3.9	585	2	US-08-426-125-2
23	91	3.9	585	2	US-08-455-355-2
24	91	3.9	1026	2	US-08-542-003-6
25	90	3.9	1026	2	US-08-322-760A-6
26	90	3.9	619	1	US-07-762-132A-2
27	90	3.9	619	1	US-08-301-722A-4

28	89.5	3.9	15281	2	US-08-471-119A-2	Sequence 2, Appl
29	89	3.8	401	2	US-08-591-079-4	Sequence 4, Appl
30	89	3.8	803	1	US-08-156-232-10	Sequence 10, Appl
31	89	3.8	803	1	US-08-304-626-10	Sequence 10, Appl
32	89	3.8	803	1	US-08-316-301A-12	Sequence 12, Appl
33	89	3.8	803	2	US-08-611-928-10	Sequence 10, Appl
34	89	3.8	803	2	US-09-173-891-10	Sequence 10, Appl
35	89	3.8	803	4	US-09-076-137-12	Sequence 12, Appl
36	89	3.8	803	5	PCT-US92-03624-12	Sequence 12, Appl
37	88.5	3.8	879	1	US-08-220-151-2	Sequence 2, Appl
38	88.5	3.8	879	1	US-08-220-151-3	Sequence 2, Appl
39	88.5	3.8	879	1	US-08-413-118-2	Sequence 2, Appl
40	88.5	3.8	879	1	US-08-413-118-3	Sequence 2, Appl
41	88.5	3.8	879	1	US-08-413-118-10	Sequence 2, Appl
42	88.5	3.8	879	3	US-08-473-446-2	Sequence 106, App
43	88.5	3.8	879	3	US-08-473-446-3	Sequence 2, Appl
44	88.5	3.8	879	3	US-08-473-446-106	Sequence 3, Appl
45	88.5	3.8	1259	3	US-09-187-049-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-07-689-008-2
Sequence 2, Application US/07689008
Patent No. 5268274
GENERAL INFORMATION:
APPLICANT: Ben-Bassat, Arie
APPLICANT: Calhoon, Roger D
APPLICANT: Fear, Anna L
APPLICANT: Gelfand, David H
APPLICANT: Meade, James H
APPLICANT: Tal, Ronny
APPLICANT: Wong, Hing
APPLICANT: Benziman, Moshe
TITLE OF INVENTION: METHODS AND NUCLEIC ACID SEQUENCES FOR THE
NUMBER OF INVENTION: EXPRESSION OF CELLULOSE SYNTHASE OPERON
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: McCutchen, Doyle, Brown & Emersen
STREET: Three Embarcadero Center
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/689,008
FILING DATE: 19910422
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 337,194
FILING DATE: 12-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 496,236
FILING DATE: 23-MAR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Murphy, Lisabeth Felix
TELEPHONE: (415) 393-2286
TELEFAX: (415) 393-2286
TELE: 340817 MACPAG SFO
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3031 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear

MOLECULE TYPE: protein
US-07-689-008-2

Query Match 4.7%; Score 110; DB 1; Length 3031;
Best Local Similarity 21.9%; Pred. No. 0.23;
Matches 79; Conservative 43; Mismatches 127; Indels 112; Gaps 14;

24 SDVSLTKEETALKES---GSNEALALV---EGIKGEIOTKPPRGQATKILGSVG 76
1423 TDESTITGVAAAMITEAESPLASGRITVALISGGGGLNLVQILAAQRKKA-KIQGDLV 1481
77 SALGKLSGATKIIISCLDIVAGIATTFGGPYMGIG-----AVASVSSITL 124
1482 LAHG---DILTSRSSPLTV-----GTVPLEPDMYMHNSRVIVGLLCIL 1529
125 SLFTGSSAKSVAAVIDRALSKHDEAIQRHAAGAKRDEAFESSAFIO---VAKQSNLT 180
1530 -----IVAVVVRALAKH---ALRR---RELEGERORTMNRVIVLSGALL 1570
181 DSDSLIAANYPVKFNSNFIGQLSRIISGAAATSLSDAKRAVDIFILYQGLVYMRFTL 240
1571 ASSCMYLVAVPVAR-----AQQASTRATTAATATRA-----PROILL 1609
241 VDLAILYKGNAEHVASAVENANRV-----NKEIADTLDLFLKLI 281
1610 QGARFWLQOQOYDANARALQNAERIANPSPDVLEVEGYOTAIGNREADTLRHLOQVA 1669
282 PEQALIGAVHPIS-----ASETSKAILNTYTFGVDPVPRPIGNRRY 324
1670 PGSAAGNLNDLSERAISQSDLSQIRSLASGONAAVAGYQKLFHGKRPISLAVEY 1729
325 K 325
DB 1730 Q 1730

RESULT 2
US-09-413-814-80
Sequence 80; Application US/09413814
Patent No. 6225064
GENERAL INFORMATION:
APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
APPLICANT: Bristol-Myers Squibb, Co.
APPLICANT: Beyer, Stefan
APPLICANT: Bloecker, Helmut
APPLICANT: Brandt, Petra
APPLICANT: Cino, Paul M
APPLICANT: Dougherty, Brian A
APPLICANT: Goldberg, Steven L
APPLICANT: Hofle, Gerhard
APPLICANT: Mueller, Joachim
APPLICANT: Reichenbach, Hans
TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
FILE REFERENCE: PCT/US 99/23535
CURRENT APPLICATION NUMBER: US/09/413, 814
EARLIER FILING DATE: 1998-10-07
EARLIER APPLICATION NUMBER: DE 198 46 493.2
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 80
LENGTH: 3079
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-413-814-80

Query Match 4.3%; Score 100.5; DB 4; Length 3079;
Best Local Similarity 22.5%; Pred. No. 2.3; Indels 113; Gaps 18;
Matches 83; Conservative 47; Mismatches 126;

38 KEASGSNEALALEGLKGEIOTKPDR-----VGQATKILGSVG 76
55 REAAGAEVYWRKALDGTATIDLPDRBARHAGARCRGAHAITLPKPLTGALARLBERG 114
77 SALGKLSGATKII---SCCLDIVAGI-----ATTFGG-----VPGSIGVAVAS 118
115 TTFESVLLSALTLLHRASQSDLVGVGPSAGRHDSEARAFGYOMLVPRVALRGAAS 174
119 FVESITSLFTGSSAKSVAAVIDRALS---KRPDEAIQRHAAGAKRDEAFESSAFIOVAKQ 176
175 FDA-----LVANRDAFLDALAHGSALRHLLARNGEAGQRALFVNAFAQ 221
177 SN---LSDSLIAANYPVKFNSNFIGQLSRIISG---AATSLSDAKRAVD---ILYQ 231
222 STPEPLDARLSALA-----ICVGDVRIAGCELETTTLADEQAAAEFDLALFAE 270
232 L---VYMR---ELLVDLAILYKGNAEH---VASAVENANRVNKEIADTLDLFLKLI 281
271 LDAGIALRFYEDQLEFPATIERM---ARHFVLLSEVSEHGRPLSELRM---LSDAERALL 327
282 -----PEQALIGAVHPISASETSKAILNTYTFGVDPVPRPIGNRRY 325
328 LDDMSGAAARQAASAPACVHAHLEFAHAAQPDATALE-----GHQRF 374
326 FTN-SYVNT 333
375 YAOIESTWT 383
DB

RESULT 3
US-08-968-685A-10
Sequence 10; Application US/08968685A
Patent No. 6214981
GENERAL INFORMATION:
APPLICANT: TUCKER, KENNETH
APPLICANT: PLOSILA, LAURA
TITLE OF INVENTION: MORAXELLA CATARRHALIS OUTER MEMBRANE
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/968, 685A
FILING DATE: No. 6214981ember 12, 1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Baldwin, Geraldine F.
REGISTRATION NUMBER: 31,232
REFERENCE/DOCKET NUMBER: 7969-060
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 2123 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-968-685A-10

Query Match 4.2% Score 97; DB 4; Length 2123;
Best Local Similarity 21.9%; Pred. No. 2.8; Mismatches 127; Indels 100; Gaps 16;
Matches 75; Conservative 40;

QY 42 GSNMALEALEGLKGLTQTPDR--VGOATKILGVSALG--KL--NSGDATKI---I 91
DB 70 GISEA-----DGGKGGANMAGDKSIAIGDIAQALGSGSIAIGDKIYHNSNNNANITAKA 124
QY 92 SGCIDYAGIATTEGPGVNGIGA-----VASFVSLSLFTGSSANNVAAVIDRA 143
DB 125 SGNESTAIIGDDVYASGASIAISDILYKRETVQOISELLPIIRGOKALNDIYQIADTN 164
QY 144 LSHHDEAIOHRAAGAKRDEAFESSAFIOVKQOONLTDLSLITANVPIYKSNFICQL 203
DB 185 LQKRRTHAQGHSTA-----VGAAMSYAKGHFSNAFGTR 218
QY 204 ESISGGAATSLSDAKRAVDLILYCOLVMEETLLVDLAILYKRGNAEHV---ASAVE 260
DB 219 ATAEGTYSIAVGLTATAKASSIAV-----GSMOATIGFAATAVG 258
QY 261 NARVVKELADTLDLHLKLIPEQALIGAVYHPIASSETSKALNTKYEGVDPVPRPIG 320
DB 259 GSQVNLNGI-ALFGSQVLOKNDVNA-----NVRAY--APDDNQPID 301
QY 321 NRRYK--FTNSYWNYSICSEAYMGNMFRGCSNVENPIRY 360
DB 302 N-RYKATFKNGATDVESI-----GN--SNGDSITRKILINV 334

RESULT 4
US-08-989-925-1

Sequence 1, Application US/08989925
Patent No. 5989820

GENERAL INFORMATION:

APPLICANT: Bandman, Olga

APPLICANT: Lal, Preeti

APPLICANT: Shah, Purvi

TITLE OF INVENTION: HUMAN ADIPOPHILIN-LIKE PROTEIN

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Inocyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/989,925

FILING DATE: Herewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0440 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 434 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:
LIBRARY: BRAITUT13
CLONE: 1620223
US-08-989-925-1

Query Match 4.1% Score 95; DB 2; Length 434;
Best Local Similarity 24.5%; Pred. No. 0.31;
Matches 62; Conservative 35; Mismatches 94; Indels 62; Gaps 10;

QY 17 SANHGR-----SDVNSLTFTVETALKKASGSNEALEALGLKGLTQTPDRVGOAT 69
DB 62 AAEKGVRTLLAAVSGAQPILSKLEPQIASASYAHRLDKLEENLPILOPTEKVLADT 121
QY 70 KILGVSALGKILNSGDATKIISGCLDIYAGIATTEGPGVNGIGAASFVSSILSLFT- 128
DB 122 KEL--VSSKV---SG-AQEMVSSAKDTV---ATQISEAVDATRGAVOSGVDTKSVYTG 171
QY 129 -----GSSAKNSVAAVIDRALSKHDEAIOHRAAGAKRDEAFESSAFIOVKQOONLTD 182
DB 172 GVSQVNGSLGQVNLGCVTLGK--SEWADNHL-----LTDA 209
QY 183 DLSITANVPIYKFSNF-----IGLESRIISGGAATTSL-----SDAKRAVDPI 226
DB 210 ELARIATSLDGFVAVSQOQROEQSYFVRLGSLSERLROHAIHSLGLKLRATKORAEAL 269
QY 227 LLYCOLVMEETL 239
DB 270 LQLSQALSIMETV 282

RESULT 5
US-08-803-973-2

Sequence 2, Application US/08803973
Patent No. 5773701

GENERAL INFORMATION:

APPLICANT: Braun, Carl J.

APPLICANT: Hemenway, Cynthia L.

APPLICANT: Turner, Nilgun E.

TITLE OF INVENTION: Plants Resistant to Infection by PVX

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF

STREET: 700 Chesterfield Village Parkway

CITY: St. Louis

STATE: Missouri

COUNTRY: USA

ZIP: 63198

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/803,973

FILING DATE: 21-FEB-1997

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/804,862

FILING DATE: 04-DEC-1991

APPLICATION NUMBER: US 07/771,912

FILING DATE: 04-OCT-1991

ATTORNEY/AGENT INFORMATION:

NAME: Hoerner Jr., Dennis R.

REGISTRATION NUMBER: 30,914

REFERENCE/DOCKET NUMBER: 38-21(10541)A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (314)537-6099

TELEFAX: (314)537-6047

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1456 amino acids

TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-803-973-2

Query Match 4.1%; Score 95; DB 1; Length 1456;
Best Local Similarity 19.0%; Pred. No. 2.4;
Matches 105; Conservative 67; Mismatches 188; Indels 194; Gaps 27;

QY 52 EELKEIQTPDRVQGAATIIISVGS-----ALGKINSG-DATKIISG---CLDIYA 99
DB 714 EYLTKIESLTERKVATCVIHAGSGSKSHAIQKALREIGKSDITVVLPTNELRLDMSK 773
QY 100 GIATP-----FGG-----PYGMGIGAVASPVSSI-LSLFTGSSAK 133
DB 774 KVPNTEPYWFKTEKALIGTGSIVIFDDYSKLPBGY-IEALICYFSKIKVLILGDSRQ 832
QY 134 NSVAVIDRALSKHDEAIQ-----RHAAGA---KRPFASSAFIOVMKQOONLTDSD 183
DB 833 SYVHETAEDASIRHLGPATEYFSKYCRYLNTATHRNKKDLAN--MLGVYSERTGVTEIS 889
QY 184 LSI-----IAANVP-----VYKFSNFIQLESRI-----SOGAA 212
DB 890 MSAEFLGPIPTLVPSDEKRLKLYMGTRNDFTYAGCGGLTKRKVOIVLDHNTQVCSANYM 949
QY 213 TSLSDAKRAVDFT-----LLYCOLVVMRETLVLDLAILYRKGAHEV 255
DB 950 YVALSRATRIHFVNTSANSASFWEKLDSTPYLKTFLSVRRQALRE---YEPAAEPI 1005
QY 256 AS-----AVENANRVKELADTLD-----FLHKLIP 282
DB 1006 REPEPOTMHCVENESVLEEYKEELLEKEPREIHSESHGSHNCVOTEDTTIOLFSHQAK 1065
QY 283 EQLLGAVYH---PISASETS-----KALINTYKFGVDPVPRIGNRRYKFT 327
DB 1066 DETLLMATIDARLKISNOTNPREPLSKDIDGVLFNTOKAMGLP-----KEPIPS 1118
QY 328 NSYWNYSICSEAYMGNYMFRGCSNVNRPNIYVSKMSDGFYTMENSDRKKLYITKHQGM 387
DB 1119 QEWMA---CAHEVQSKYLSKCNLINCTVROS-----PDEDKIMVFLKSQ-- 1164
QY 388 GNGTLEDGDDGCHRAFIP-LRHGKTMV-----SKRWPMWFMYMESSASGYR 435
DB 1165 -WYTKYEKLG-----LPKIRPGOTIAAFYQOTVLMFGTMARYMWFQAFQPKKEVFI- 1215
QY 436 SWENNPGPOGHMSI 449
DB 1216 NCETTPEDMSANAL 1229

RESULT 6
US-08-803-972-2
Sequence 2, Application US/08803972
Patent No. 5792937
GENERAL INFORMATION:
APPLICANT: Braun, Carl J.
APPLICANT: Hemenway, Cynthia L.
APPLICANT: Turner, Nilgun E.
TITLE OF INVENTION: Plants Resistant to Infection by PVX
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. B44F
STREET: 700 Chesterfield Village Parkway
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/803, 972
FILING DATE: 21-FEB-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/804, 862
FILING DATE: 04-DEC-1991
APPLICATION NUMBER: US 07/771, 912
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.
REGISTRATION NUMBER: 30,914
REFERENCE/DOCKET NUMBER: 38-21(10541)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)537-6099
TELEFAX: (314)537-6047
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1456 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-803-972-2

Query Match 4.1%; Score 95; DB 1; Length 1456;
Best Local Similarity 19.0%; Pred. No. 2.4;
Matches 105; Conservative 67; Mismatches 188; Indels 194; Gaps 27;

QY 52 EELKEIQTPDRVQGAATIIISVGS-----ALGKINSG-DATKIISG---CLDIYA 99
DB 714 EYLTKIESLTERKVATCVIHAGSGSKSHAIQKALREIGKSDITVVLPTNELRLDMSK 773
QY 100 GIATP-----FGG-----PYGMGIGAVASPVSSI-LSLFTGSSAK 133
DB 774 KVPNTEPYWFKTEKALIGTGSIVIFDDYSKLPBGY-IEALICYFSKIKVLILGDSRQ 832
QY 134 NSVAVIDRALSKHDEAIQ-----RHAAGA---KRPFASSAFIOVMKQOONLTDSD 183
DB 833 SYVHETAEDASIRHLGPATEYFSKYCRYLNTATHRNKKDLAN--MLGVYSERTGVTEIS 889
QY 184 LSI-----IAANVP-----VYKFSNFIQLESRI-----SOGAA 212
DB 890 MSAEFLGPIPTLVPSDEKRLKLYMGTRNDFTYAGCGGLTKRKVOIVLDHNTQVCSANYM 949
QY 213 TSLSDAKRAVDFT-----LLYCOLVVMRETLVLDLAILYRKGAHEV 255
DB 950 YVALSRATRIHFVNTSANSASFWEKLDSTPYLKTFLSVRRQALRE---YEPAAEPI 1005
QY 256 AS-----AVENANRVKELADTLD-----FLHKLIP 282
DB 1006 REPEPOTMHCVENESVLEEYKEELLEKEPREIHSESHGSHNCVOTEDTTIOLFSHQAK 1065
QY 283 EQLLGAVYH---PISASETS-----KALINTYKFGVDPVPRIGNRRYKFT 327
DB 1066 DETLLMATIDARLKISNOTNPREPLSKDIDGVLFNTOKAMGLP-----KEPIPS 1118
QY 328 NSYWNYSICSEAYMGNYMFRGCSNVNRPNIYVSKMSDGFYTMENSDRKKLYITKHQGM 387
DB 1119 QEWMA---CAHEVQSKYLSKCNLINCTVROS-----PDEDKIMVFLKSQ-- 1164
QY 388 GNGTLEDGDDGCHRAFIP-LRHGKTMV-----SKRWPMWFMYMESSASGYR 435
DB 1165 -WYTKYEKLG-----LPKIRPGOTIAAFYQOTVLMFGTMARYMWFQAFQPKKEVFI- 1215
QY 436 SWENNPGPOGHMSI 449
DB 1216 NCETTPEDMSANAL 1229

RESULT 7
US-08-803-972-2
Sequence 2, Application US/08803972
Patent No. 5183745-2
APPLICANT: DANCHIN, ANTOINE;GLASER, PHILIPPE;KRIN, EVELYN;

BARZU, OCTAVIEN; LADANT, DANIEL; ULLMAN, AGNES
 ; BIOLOGICAL USES
 ; NUMBER OF SEQUENCES: 13
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/426,541
 ; FILING DATE: 25-OCT-1989
 ; SEQ ID NO: 2
 ; LENGTH: 1489
 ; 5183745-2

Query Match 4.1%; Score 94.5; DB 6; Length 1489;
 Best Local Similarity 25.0%; Pred. No. 2.8;
 Matches 65; Conservative 49; Mismatches 115; Indels 31; Gaps 12;

OY 11 IYLAITSARKHRSDVNSLITKVEETALKESGSENEALEALGLKEIO-----TKPDYV 65
 DB 587 IALALAAR-----GVTSGLVAGASAGAAAGALAAALSPME-IYGLVQOSHAYADQDLKL 640
 OY 66 GOATKILGVSALGKLNAGDATKIISGLDIYAGIATTFPGPVGMGIGA--VASEVSSI 123
 DB 641 AQESSAYGEEDALLAOLYRDKT-AAEGAVAGSAVISTVGAASVIAAASVGAAPAVY 699
 OY 124 LSFTGSSAKNSVAVIDRALSKH--RDEAIORHAAGAKRDEFAESSAFIQVMQOQSLTD 181
 DB 700 TSLTGT--ALNGILRGVQOPIETKLANDYARKIDELGGPAYPEKN--LQARHEOLANSND 755
 OY 182 SLSITIAANPVYKFSNFIOLESRIISOG-----AATTSLSDAKRAVD-FILLYCO-LVYM 235
 DB 756 GLRKMILADLOAGWNASVIGVOTTEISKALIELAATGNADNLKSVDFVDFVQGERVA 815
 OY 236 RETLLVDA-----ILYRKG 250
 DB 816 GQPVLDVAAGIDIASRKG 835

RESULT 8
 5183745-6
 ; Patent No. 5183745
 ; APPLICANT: DANCHIN, ANTOINE; GLASER, PHILIPPE; KRIN, EVELYN;
 ; BARZU, OCTAVIEN; LADANT, DANIEL; ULLMAN, AGNES
 ; TITLE OF INVENTION: ADENYL CYCLASE DERIVATIVES AND THEIR
 ; BIOLOGICAL USES
 ; NUMBER OF SEQUENCES: 13
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/426,541
 ; FILING DATE: 25-OCT-1989
 ; SEQ ID NO: 6
 ; LENGTH: 1794
 ; 5183745-6

Query Match 4.1%; Score 94.5; DB 6; Length 1794;
 Best Local Similarity 25.0%; Pred. No. 3.9;
 Matches 65; Conservative 49; Mismatches 115; Indels 31; Gaps 12;

OY 11 IYLAITSARKHRSDVNSLITKVEETALKESGSENEALEALGLKEIO-----TKPDYV 65
 DB 586 IALALAAR-----GVTSGLVAGASAGAAAGALAAALSPME-IYGLVQOSHAYADQDLKL 639
 OY 66 GOATKILGVSALGKLNAGDATKIISGLDIYAGIATTFPGPVGMGIGA--VASEVSSI 123
 DB 640 AQESSAYGEEDALLAOLYRDKT-AAEGAVAGSAVISTVGAASVIAAASVGAAPAVY 698
 OY 124 LSFTGSSAKNSVAVIDRALSKH--RDEAIORHAAGAKRDEFAESSAFIQVMQOQSLTD 181
 DB 699 TSLTGT--ALNGILRGVQOPIETKLANDYARKIDELGGPAYPEKN--LQARHEOLANSND 754
 OY 182 SLSITIAANPVYKFSNFIOLESRIISOG-----AATTSLSDAKRAVD-FILLYCO-LVYM 235
 DB 755 GLRKMILADLOAGWNASVIGVOTTEISKALIELAATGNADNLKSVDFVDFVQGERVA 814

OY 236 RETLLVDA-----ILYRKG 250
 DB 815 GQPVLDVAAGIDIASRKG 834

RESULT 9
 US-08-591-079-2

; Sequence 2, Application US/08591079
 ; Patent No. 5972899

; GENERAL INFORMATION:
 ; APPLICANT: Zychlinsky, Arturo

; TITLE OF INVENTION: Apoptosis Induced by Shigella Ipab

; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:

; ADDRESS: MORRISON & FOERSTER
 ; STREET: 2000 Pennsylvania Avenue, NW

; CITY: Washington
 ; STATE: DC

; COUNTRY: USA
 ; ZIP: 20006-1812

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/591,079
 ; FILING DATE:

; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:

; NAME: Livnat, Shmuel
 ; REGISTRATION NUMBER: 33,949

; REFERENCE/DOCKET NUMBER: 15661-20017.00
 ; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 887-0764
 ; TELEFAX: (202) 887-0764

; TELETYPE: 90-4030 MRSNORERSWSH
 ; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 580 amino acids

; TYPE: amino acid
 ; TOPOLOGY: linear

; MOLECULE TYPE: protein
 ; US-08-591-079-2

Query Match 4.0%; Score 93.5; DB 2; Length 580;
 Best Local Similarity 22.1%; Pred. No. 0.73;
 Matches 68; Conservative 47; Mismatches 114; Indels 79; Gaps 14;

OY 39 EASGSENEAL--EALGLKEIOTPD-----RVGQATKILGVSALGKLNAGDATK 89
 DB 265 EESLKNLADLFQSLQESKTEMERKSDYAAEVRAEELNRYMGCVKILGL----- 317
 OY 90 IISGLDIYAGIATTFPGPVGMGIGANA--SFVSSISLFTGSS-----ANNSVAAYDR 142
 DB 318 -----LITVSVAFAAFSGGASIALAAGLAWTDALYOAATGNSFMEQALNPIKMAVIE 372
 OY 143 ALSKRDDEAIORHAAGAKRDEFAESSAFIQVMQOQSLTDLSITIAANPVYKFSNFIO 202
 DB 373 PLIKLISDATTKMEIGLVGDSKRAKMGISIL--GAIAGALVLYAAVLYAT-----VGK 424
 OY 203 LESRISQGAATTSLSDAKRAVD-FILLYCOLVYMRRELLVNDLALLYRKGNAEHVASAENA 262
 DB 425 -----QAALAEINIGK-----IIGKT-LTDLIPKFLKNFSQDLDTLITNA 464
 OY 263 -NRVNEKLAADLDLFLKLI-----PEQALIG-----AYYHPIASSETSK 301
 DB 465 VARLKKFTLAGADEVYSKQIISTHLNQAVLLGESVNSATQAGGSVASAVFO--NSASTML 522
 OY 302 ALINTYTK 309

Db 523 ADLFLSKY 530

RESULT 10
US-08-480-604A-28Sequence 28, Application US/08480604A
Patent No. 5736139

GENERAL INFORMATION:

APPLICANT: KINK, JOHN A.

APPLICANT: THALLEY, BRUCE S.

APPLICANT: PADHAY, NISHA V.

APPLICANT: FIRCA, JOSEPH R.

APPLICANT: STAFFORD, DOUGLAS C.

TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND PREVENTION OF C. DIFFICILE DISEASE

NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:

ADDRESSEE: MEDLEN & CARROLL, LLP

STREET: 220 MONTGOMERY STREET, SUITE 2200

CITY: SAN FRANCISCO

STATE: CALIFORNIA

COUNTRY: UNITED STATES OF AMERICA

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION NUMBER: US/08/480,604A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/422,711

FILING DATE: 14-APR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/405,496

FILING DATE: 16-MAR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/329,154

FILING DATE: 25-OCT-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/161,907

FILING DATE: 02-DEC-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/985,321

FILING DATE: 04-DEC-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/429,791

FILING DATE: 31-OCT-1989

ATTORNEY/AGENT INFORMATION:

NAME: INGOLIA, DIANE E.

REGISTRATION NUMBER: 40,027

REFERENCE/DOCKET NUMBER: OPND-01763

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 705-8410

TELEFAX: (415) 397-8338

INFORMATION FOR SEQ ID NO: 28:

SEQUENCE CHARACTERISTICS:

LENGTH: 1296 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-480-604A-28

Query Match 4.0%; Score 93; DB 1; Length 1296;

Best Local Similarity 20.9%; Pred. No. 3.2;

Matches 112; Conservative 61; Mismatches 161; Indels 202; Gaps 28;

Db 9 LFVIALTSKRGK-----RSDVNSL---TKVET-----ALKASGSNEA--LEALEG 53
550 MFHILRAQEFHSGSRIALFNSVNEALINPSRYVTFSSDYKVKVNALEAFLGNVEQ 609

QY 54 L-----KGEIQKPDVGOATKILGVSALGKLN SGD-----ATKIISGC---L 95

Db 610 LVYDFTDETSEVST-VDKINDITIIIPYIGPA---LNIIGNMLYKDEFGALIFSGAVILL 665

QY 96 DIVAGIATTFGPGVGMGIGAVASVSSILSLFTGSSAKNSVAAV--IDRALSKHRD---- 149

Db 666 EFPEIAI---PV-LGTFALVSYIA-----NKVLTVQTDINALSKNEKWE 708

QY 150 -----EAIORHAAGAKRDF-AESSAFIQVMKQSNLSDSD 183

Db 709 VYKVIYVNWMLAKVNTQIDLIKKKKEALENQAETKALINQVNOYQTEEEKNINIFNID 768

QY 184 L-----SIIANPVYKFSNFIG---QLESRIQGAATSLSDAKRAVDLFLYCOLV 234

Db 769 LSKLINESINAKMININKFNOCSSVYLSNMIPIYG-----VKRIEDF----- 811

QY 235 MREPLVDLAILYRKGAHEHVASAVEN-ANRNKRELAADTDLFLKHLIPEQALIGAVHP 293

Db 812 --DASLKDALIKYIYDNGTLIGQVDRDKDKVNTLSTDIPQLSKYVNOHLLS----- 864

QY 294 ISASETSKAILNYTFYGVDPVPRPIGNRRYKFTNSYMWTVSICSEAYMGNTFRGCSNV 353

Db 865 -TFTEYIKNLIIN-----TSILNLYRESNHL----- 888

QY 354 RNPNIKVSMSDGFYTMENSDDRKLITKHDGCGWGTLDDEPDGOGHRRFIPLRHGXVM 413

Db 889 ----IDLK-----YASKINIG---SKVNFPIDKNOQLEFNLESSEKIE 925

QY 414 VSKRWPNW-FMYMESSAGYIR-----SWENPPOGSHSIT 450

Db 926 VILKNAIVNSMYENFSTFWIRIPYFNISILNNEYTIINCENNSG-----WKVS 977

RESULT 11
US-08-405-496A-28

Sequence 28, Application US/08405496A

Patent No. 5919665

GENERAL INFORMATION:

APPLICANT: WILLIAMS, JAMES A.

TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM

NUMBER OF SEQUENCES: 30

CORRESPONDENCE ADDRESS:

ADDRESSEE: MEDLEN & CARROLL, LLP

STREET: 220 MONTGOMERY STREET, SUITE 2200

CITY: SAN FRANCISCO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION NUMBER: US/08/405,496A

APPLICATION NUMBER: US/08/405,496A

FILING DATE: 16-MAR-1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/329,154

FILING DATE: 25-OCT-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/161,907

FILING DATE: 02-DEC-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/985,321

FILING DATE: 04-DEC-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/429,791

FILING DATE: 31-OCT-1989

ATTORNEY/AGENT INFORMATION:

NAME: INGOLIA, DIANE E.

REGISTRATION NUMBER: 40,027
 REFERENCE/DOCKET NUMBER: OPHD-01308
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 705-8410
 TELEFAX: (415) 397-8338
 INFORMATION FOR SEQ ID NO: 28:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1296 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-405-496A-28

Query Match 4.0%; Score 93; DB 2; Length 1296;
 Best Local Similarity 20.9%; Pred. No. 3.2;
 Matches 112; Conservative 61; Mismatches 161; Indels 202; Gaps 28;

QY 9 LEIVLAITSKHKR-----RSDVNSLL---TKVET-----ALKEASGSNEAA--LEALEG 53
 DB 550 MFHYLAQEFHKSRIALNTNSVNEALLNPSRYTTFSSDYKVKVAKTEAMFLGVEQ 609
 QY 54 L-----KGEIQTKPRVGOATKILGVSALGLNSGD-----ATKIISGC---L 95
 DB 610 LVYDFDETSEVST-TDKIADITIIIPYIGPA---LNIGNMLYKDDFVGLIFSGAVILL 665
 QY 96 DIVAGIATTFGGPGVGMIGAVASFVSSILSLFTGSSAKNSVAAY--IDRALSKHRD--- 149
 DB 666 EFIEPEIAI---PV-LGTFALVSYIA-----NKVLTVQTIIDNALSKRNEKWD 708
 QY 150 -----EAIORHAAGAKRDF-AESSAFIQVWKOQSNLTDSD 183
 DB 709 VYKIYTNWLAKVNTQIDILRKMKKEALNOAEATKAIINYQINQYTEEEKNNININIDD 768
 QY 184 L-----SITAAVVPYKFSNFTG---QLESRIISOGAATTSLSDAKRAVDFTILLYCOLVY 234
 DB 769 LSSKLINESINKAMININIKFNOCSSVSLNMSMIPYG-----VKRLDF----- 811
 QY 235 MRETLVYDALILYRKGNAEHVASAVEN-ANKVKELAADTLDLFLKLIPEQALIGAVYHP 293
 DB 812 --DASLKDALIKYTYDRGTLIGQVDRLKDKVNNLTSTDIPOLSKYVDNQRLIS----- 864
 QY 294 ISASETSKALINTKYKFGVDPVRPIGNRRYKFTNSWNTYNTYSEAYMGNYMERGCSNV 353
 DB 865 -TFETIKINII-----TSLINLRYESNHL----- 888
 QY 354 RNPNTVSKMSDGFYTWMSNDRRLKYITKHOGMGWGLTDEPDGQGHMRFILRHGKYM 413
 DB 889 ----IDLST-----YASKINIG---SKVNFDPIDKNQIOLFPLESSKIE 925
 QY 414 VSKRMPNM-FWYMESSASGYIR-----SWENNPQGHWSIT 450
 DB 926 VILKNAIVNMYENSTSEFWIRIPRYFNSISLNNETIINCENNSG---WKVS 977

RESULT 12
 US-08-915-136-28
 Sequence 28, Application US/08915136
 Patent No. 6290960
 GENERAL INFORMATION:
 APPLICANT: KINK, JOHN A.
 APPLICANT: THALLEY, BRUCE S.
 APPLICANT: PADHYE, NISHA V.
 APPLICANT: FIRCA, JOSEPH R.
 APPLICANT: STAFFORD, DOUGLAS C.
 TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
 NUMBER OF SEQUENCES: 32
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MEDLEN & CARROLL, LLP
 STREET: 220 MONTGOMERY STREET, SUITE 2200
 CITY: SAN FRANCISCO
 STATE: CALIFORNIA

COUNTRY: UNITED STATES OF AMERICA
 ZIP: 94104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC Compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/915,136
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/480,604
 FILING DATE:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/405,496
 FILING DATE: 16-MAR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/329,154
 FILING DATE: 25-OCT-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/161,907
 FILING DATE: 02-DEC-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/985,321
 FILING DATE: 04-DEC-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/429,791
 FILING DATE: 31-OCT-1989
 ATTORNEY/AGENT INFORMATION:
 NAME: INGOLTA, DIANE E.
 REGISTRATION NUMBER: 40,027
 REFERENCE/DOCKET NUMBER: OPHD-01763
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 705-8410
 TELEFAX: (415) 397-8338
 INFORMATION FOR SEQ ID NO: 28:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1296 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-915-136-28

Query Match 4.0%; Score 93; DB 4; Length 1296;
 Best Local Similarity 20.9%; Pred. No. 3.2;
 Matches 112; Conservative 61; Mismatches 161; Indels 202; Gaps 28;

QY 9 LEIVLAITSKHKR-----RSDVNSLL---TKVET-----ALKEASGSNEAA--LEALEG 53
 DB 550 MFHYLAQEFHKSRIALNTNSVNEALLNPSRYTTFSSDYKVKVAKTEAMFLGVEQ 609
 QY 54 L-----KGEIQTKPRVGOATKILGVSALGLNSGD-----ATKIISGC---L 95
 DB 610 LVYDFDETSEVST-TDKIADITIIIPYIGPA---LNIGNMLYKDDFVGLIFSGAVILL 665
 QY 96 DIVAGIATTFGGPGVGMIGAVASFVSSILSLFTGSSAKNSVAAY--IDRALSKHRD--- 149
 DB 666 EFIEPEIAI---PV-LGTFALVSYIA-----NKVLTVQTIIDNALSKRNEKWD 708
 QY 150 -----EAIORHAAGAKRDF-AESSAFIQVWKOQSNLTDSD 183
 DB 709 VYKIYTNWLAKVNTQIDILRKMKKEALNOAEATKAIINYQINQYTEEEKNNININIDD 768
 QY 184 L-----SITAAVVPYKFSNFTG---QLESRIISOGAATTSLSDAKRAVDFTILLYCOLVY 234
 DB 769 LSSKLINESINKAMININIKFNOCSSVSLNMSMIPYG-----VKRLDF----- 811
 QY 235 MRETLVYDALILYRKGNAEHVASAVEN-ANKVKELAADTLDLFLKLIPEQALIGAVYHP 293
 DB 812 --DASLKDALIKYTYDRGTLIGQVDRLKDKVNNLTSTDIPOLSKYVDNQRLIS----- 864

OY 294 ISASETSKALINRYKGVDPVPPRIGNRRYKFTNSWNTYSIGSEAYMGVYMRGCSNV 353
: : : : :
Db 865 -TFEYIKININ-----TSIINLRYESNHL----- 888
OY 354 RNPRIYSKMSDGFYTMENSRRRLYITKHQGWGWTGTDPPDQGMRIPIRLHGYM 413
: : : : :
Db 889 ----IDSR-----YASKINIG---SKVNDPIDKNOIOIFNESSKIE 925
OY 414 VSSKRWPMW-FMYMESSAGYIR-----SWENNPQPGQMSIT 450
: : : : :
Db 926 VILKNALVYNSMVENFSTSEWIRIPKYFNSTLNNETIINCMENNSG----WKVS 977

RESULT 13

US-08-705-660-18
; Sequence 18, Application US/08705660
; Patent No. 585683
; GENERAL INFORMATION:
; APPLICANT: KEESEE, SUSAN
; APPLICANT: OBAR, ROBERT
; APPLICANT: WU, YING-YJE
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; TITLE OF INVENTION: DETECTION OF CERVICAL CANCER
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault
; STREET: 125 High St.
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/705,660
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: GREENHALGH, DUNCAN A
; REGISTRATION NUMBER: 38,678
; REFERENCE/DOCKET NUMBER: MTP-023 (8395/27)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-705-660-18

Query Match

Best Local Similarity 4.0%; Score 92.5; DB 2; Length 432;
Matches 64; Conservative 44; Mismatches 117; Indels 43; Gaps 13;

OY 35 TALKEAGSNEALEALEGLKGEIOTKPDVVGQATKILGSVSGALGKLNDSGATKIISGC 94
: : : : :
Db 12 SSISGSSGLGGSSRTSCRLSGGLGAGSCRLGSA---GGIGSTLGGSSYSSCYSPGSG- 66
OY 95 LDIYAGIATTFGPGVGMGIGAVASFVSSILSLFTGSSAKNSVAVID--PALSXHRDEA- 151
: : : : :
Db 67 -----GGYSSFGVDGLAGGEKA-----TMQNLNRLASYLDKVRALEBANTELE 113
OY 152 -----IQRAHAGAKRPFASFAFIQVMKQOOSNLTDSD-----LSITANVPYKFSNF 199
: : : : :
Db 114 VKIRDWYQRCAPGPARPARYSOYRTIELOKILATVDNANILQIDNARLAADDFRTKF 173

OY 200 IGOLESIRISOGAATTSLSDAKRAVD-FILLYCOLVVRRETLVDLATILYRKGNAEHVASA 258
: : : : :
Db 174 ETEQALRLSVEA---DINGLRVYDELTLARADEMQIENLKEELA--YLKKNHEEMANA 226
OY 259 VEN--ANRVKEL-AADTLDFLHKLPIE 283
: : : : :
Db 229 LRQGVGEINVEDMADGVD-LSRLINE 255

RESULT 14

US-08-989-045-18
; Sequence 18, Application US/08989045
; Patent No. 6027905
; GENERAL INFORMATION:
; APPLICANT: KEESEE, SUSAN
; APPLICANT: OBAR, ROBERT
; APPLICANT: WU, YING-YJE
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; TITLE OF INVENTION: DETECTION OF CERVICAL CANCER
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault
; STREET: 125 High St.
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/989,045
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: GREENHALGH, DUNCAN A
; REGISTRATION NUMBER: 38,678
; REFERENCE/DOCKET NUMBER: MTP-023 (8395/27)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-989-045-18

Query Match

Best Local Similarity 4.0%; Score 92.5; DB 3; Length 432;
Matches 64; Conservative 44; Mismatches 117; Indels 43; Gaps 13;

OY 35 TALKEAGSNEALEALEGLKGEIOTKPDVVGQATKILGSVSGALGKLNDSGATKIISGC 94
: : : : :
Db 12 SSISGSSGLGGSSRTSCRLSGGLGAGSCRLGSA---GGIGSTLGGSSYSSCYSPGSG- 66
OY 95 LDIYAGIATTFGPGVGMGIGAVASFVSSILSLFTGSSAKNSVAVID--PALSXHRDEA- 151
: : : : :
Db 67 -----GGYSSFGVDGLAGGEKA-----TMQNLNRLASYLDKVRALEBANTELE 113
OY 152 -----IQRAHAGAKRPFASFAFIQVMKQOOSNLTDSD-----LSITANVPYKFSNF 199
: : : : :
Db 114 VKIRDWYQRCAPGPARPARYSOYRTIELOKILATVDNANILQIDNARLAADDFRTKF 173
OY 200 IGOLESIRISOGAATTSLSDAKRAVD-FILLYCOLVVRRETLVDLATILYRKGNAEHVASA 258
: : : : :
Db 174 ETEQALRLSVEA---DINGLRVYDELTLARADEMQIENLKEELA--YLKKNHEEMANA 226

OY 259 VEN--ANRVNKL--ADTLDFLHKLPE 283
Db 229 LRGVGGFINVMDAPGVD--LSRLINE 255

RESULT 15

US-08-694-865-16
Sequence 16, Application US/08694865
Patent No. 5837268

GENERAL INFORMATION:

APPLICANT: POTTER, ANDREW A.
APPLICANT: MANN, JOHN G.
TITLE OF INVENTION: GHRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:

ADDRESSEE: REED & ROBINS LLP
STREET: 285 HAMILTON AVENUE, SUITE 200
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94301

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/694,865
FILING DATE: 09-AUG-1996
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)327-3400
TELEFAX: (415)327-3231

INFORMATION FOR SEQ. ID NO: 16:

SEQUENCE CHARACTERISTICS:
LENGTH: 699 amino acids
TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-694-865-16

Query Match 4.0%; Score 92.5; DB 2; Length 699;
Best Local Similarity 18.6%; Pred. No. 1.3;

Matches 93; Conservative 61; Mismatches 200; Indels 147; Gaps 21;

OY 24 SPVNSLITVY-----ETALKRSQSGSENALEALBGLKGEIQTTPDVGQATKILGSVGS 77
Db 219 SGISQILGSLVLAGMDLDELALQNNNSNOHALAKAGLELTNSLI-----ENIANSVKTIDEPGE 274
OY 78 ALGKINS-----GDATKTIISGC-----LDIVAGIAT-----TFGG 107
Db 275 QISQFGSKLQNKIGLGTLDKDKLNIGLDKAGLGDVLSGLSGATAALYLVADKNASTAK 334
OY 108 PYGNG-----IGAVASFVSSIL---SLFTSSAKNSVAAYIDRLSK----- 146
Db 335 KYGAGFELANGVYGNITKAVSSYIIAQVYAGLSSTGPAAIIASTVSLAISPLAPAGIA 394
OY 147 ---HDEDAIORHAGAKDFPSSAFIOVMKOOSNLTDSDLSIIANVPVYKFSNFIGOL 203
Db 395 DKFNNAKSLSEYAEFFKRLGYDGNMLAEYOGTGTIDASYTAI-----NTALAAI 445
OY 204 ESRISQGAATSLSDAKRAVPFILLYCOLV-----MRETLVDLALILYRKG 250
Db 446 AGGVSAALADLTPEVKVH-----NLVITNSKKREKVTIOMFREADFAKEVPNYKAT 496
OY 251 NAEHVASAV--ENANRVNKLADTLDFLHKLIPGQALIGAVYHPISASETSKAILNNTKY 309
Db 497 KDEKIEETIIGONGERITSKYQVD-----LIAKNGKITDDELKSVYDNT--- 540

OY 310 FGVPDVPPIGNRRYKFTNSYWNVTYSICEAYMGNYMRGCSNVRNPINRYSKMSDGEFT 369
Db 541 -----ELKHKNVTNNSLDKLISVS-----AFTSSNDSRNVLVAPTSMID----- 581
OY 370 MENDRRRLYITTKHDQGWGNGTLDDEPDGDGHMRFIPLRHGKYVWSKRPMPWFMYMSS 429
Db 582 ---QSLSSLOFARGSQHWSYGL---RPGSGQDMSYGLRPGG---SSQHW-----SYGLRP 628
OY 430 ASGYIRSWEN--NPGPGCHWS 448
Db 629 GSG--SQDMSYGLRPGGSQHWS 648

Search completed: May 8, 2002, 12:13:31
Job time: 124 sec

[illegible]

xx Claim 4; Page 26-27; 32pp; Japanese.

xx This invention describes a novel protein which has hemolytic activity,

xx blood platelet agglutination activity and a molecular weight of about

xx 50,000 Da (by SDS-PAGE). The protein, homologs and analogs of the

xx invention can be used as a novel approach to developing drugs useful for

xx the treatment of jelly fish stings, pharmaceuticals with blood platelet

xx agglutination activity, pesticides by use of the hemolytic activity, and

xx in the study of the hemolytic mechanism. This sequence represents the

xx hemolytic protein described in the invention

xx

xx Sequence 450 AA;

xx

xx Query Match 100.0%; Score 2322; DB 20; Length 450;

xx Best Local Similarity 100.0%; Pred. No. 1.2e-195;

xx Matches 450; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

xx

QY 1 MILKLPWLFYLAITSKAKHGKRSQVNSLTFKVFALKEASGSNEALEGLKEGIOT 60

DB 1 mlklhlpwlfyaltstakbhgrsvnslltkvetalkaagsneaaleglkgeigt 60

QY 61 KPDVGAQAKIIGSVGALGKINSDAQKIIISGLDIVAGIATTFGPGVGMGICAVASV 120

DB 61 kpdvgaqatkligsvsaglkingsdackllsgcidlvagiatftgpgvmgigavasfv 120

QY 121 SSIISLFGSSAKNSVAVIDRALSKHDEAIOHRAAGAKRPAESSAFIOVMKQOSULT 180

DB 121 ssiislfsgssaknsvaavidralskhdeaiorhaagakrpfassafiovmkqosult 180

QY 181 DSLSIIAANVPYKFSNFICGLESRIISOGAATSLSDAKRAVDIILLYCOLVWRETL 240

DB 181 dsdlsiiaanvpkyfnsnficglesiisogaatsslsdakravdiillycolvwretil 240

QY 241 VDLALIRKGNAEHVASAVENANRYNKLADTDLFLHKLIPBQALLIGAVHPISASETS 300

DB 241 vdlalirkgnaehvasavenanrynkladtdlflhklipbqalligavhpiasets 300

QY 301 KILWTVYFEGVDPDPRPIGNRRYKFTNSYMTYSIGSFAYGNVMPFGSCSVRNPNRV 360

DB 301 kilwtyvfegvdpdprpignrryktfnsymtysicgsfaygnvmpfgscsvrnpnriv 360

QY 361 SKMSDFYTMENSDRRKLYITTKHDOGMCGLTDEDPDQGHMRFPLRHGKYMVSSKRW 420

DB 361 skmsdfytmensdrklyittkhdogmcgltdeppdqghmrfplrhgkymvsskrw 420

QY 421 NMFYTMESSASGYIRSWENNPQPGHWSIT 450

DB 421 nmfymessasgyirswennpqpghwsit 450

xx

xx RESULT 2

xx AAB13634

xx ID AAB13634 standard; Protein: 964 AA.

xx AC AAB13634;

xx DT 02-FEB-2001 (first entry)

xx DE C. trachomatis pmpe gene protein.

xx KM Chlamydial infection; sexually transmitted disease;

xx KM pelvic inflammatory disease; PID; tubal obstruction; infertility;

xx KM trachoma; blindness; acute respiratory tract infection;

xx KM atherosclerosis; coronary heart disease; antibacterial.

xx OS Chlamydia trachomatis.

xx PN WO200034483-A2.

xx PD 15-JUN-2000.

PF 08-DEC-1999; 99WO-US29012.

XX

PR 08-DEC-1998; 98US-0208277.

PR 08-APR-1999; 99US-0288594.

PR 01-OCT-1999; 99US-0410568.

PR 22-OCT-1999; 99US-0426571.

PA (CORI-) CORIAX CORP.

XX

PI Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;

XX WPI; 2000-431303/37.

DR

XX

PT Isolated polypeptide for diagnosis and treatment of Chlamydia infection

PT comprises immunogenic portion of Chlamydia antigen, which comprises

PT amino acid sequence encoded by polynucleotide sequence

PS Claim 2; Pages 184-186; 256pp; English.

XX

XX The present invention relates to new nucleic acid sequences and the

XX proteins encoded by the nucleic acid sequences. The encoded proteins

XX comprise an immunogenic portion of a Chlamydia antigen. The encoded

XX proteins are useful for the serodiagnosis and treatment of Chlamydia

XX infection. Chlamydiae are intracellular bacterial pathogens that are

XX responsible for a wide variety of human infections. C. trachomatis

XX infection is one of the most common sexually transmitted diseases and can

XX lead to pelvic inflammatory disease (PID), resulting in tubal obstruction

XX and infertility. Trachoma due to ocular infection with C. trachomatis is

XX the leading cause of preventable blindness worldwide. C. pneumonia is a

XX major cause of acute respiratory tract infections in humans and is also

XX thought to play a role in the pathogenesis of atherosclerosis and

XX coronary heart disease. The present sequence is a protein isolated in the

XX present invention.

xx

xx Sequence 964 AA;

xx

xx Query Match 4.8%; Score 112.5; DB 21; Length 964;

xx Best Local Similarity 21.3%; Pred. No. 0.66;

xx Matches 110; Conservative 61; Mismatches 184; Indels 161; Gaps 28;

xx

QY 21 GKRSNV-----NSLTKVTALKKESGSNEALEGLKEGIQKPDVGAQAKIIGSVG 76

DB 21 gkrsv-----nsltkvtalkkesgsnealeglkegiqkpdvgaqakligsvg 76

QY 77 SALCKLNSGDATFKIISGLDIVAGIATTFGPGVGMGICAVASFSIISLFGSSAKNSV 136

DB 77 salcklnsgdattkiiisgldivagiatftgpgvgmigicavasssiislfsgssaknsv 136

QY 302 -----dltknysgai-----yapvvlidvngpryfinniamkg 336

DB 302 -----dltknysgai-----yapvvlidvngpryfinniamkg 336

QY 137 AAVIDRALSKHDEAIOHRA-----AGAKR-----DFAESSAFIOV 172

DB 137 aavidralskhdeaiorha-----agakr-----dfaessafioy 172

QY 337 gelyldgtsnakiad-dthalifoenlvtnanglatsnprnaltvaassgeall 395

DB 337 gelyldgtsnakiad-dthalifoenlvtnanglatsnprnaltvaassgeall 395

QY 173 MKQOS-NLTDSD-ISIIAANVPYKFSNFICGLESRIISOGAATSLSDAKRAVD----- 224

DB 173 mkqos-nltdsd-isiiiaanvpkyfnsnficglesiisogaatsslsdakravd----- 224

QY 396 gagsqnliflydplevsnagvsv-sfkkeadqtsvsvtsadfhqnlqtkcpap 454

DB 396 gagsqnliflydplevsnagvsv-sfkkeadqtsvsvtsadfhqnlqtkcpap 454

QY 225 -----FILL--YCOLVWRETL--LVDL--ATL--YRKGNAEHVASAVENANRYNKE 268

DB 225 -----fill--ycolvwretil--lvdl--atl--yrkgnaehvasavenanrynke 268

QY 455 ltlngficiedhaglvtnrftqggyvsjngavlsckvngtd-----sasnasitlkh 510

DB 455 ltlngficiedhaglvtnrftqggyvsjngavlsckvngtd-----sasnasitlkh 510

QY 269 LAADTLDFLHKLIPBQALLIGAVHPISASSETSKALINTY-----KYGVDPDPRPI-----G 320

DB 269 laadtldflhklipbqalligavhpiasetskailnty-----kygvdpdprpi-----g 320

QY 511 lgl-----lssllksgaei-----plwveplnsmnyladtaafsisdvlsiddyg 561

DB 511 lgl-----lssllksgaei-----plwveplnsmnyladtaafsisdvlsiddyg 561

QY 321 NRRKFTNSYNTYSIGCEAFAMGNVMPFGSCSVRNPNRVKMSMGFTMTMSDRKLYI 380

DB 321 nrrkftnsyntysigceafamgnvmpfgscsvrnpnrvmksmgftmtmsdrklyi 380

QY 562 nspyseldi-----thalsq-----pmlsiseasdnqjseidfsqjnv 602

DB 562 nspyseldi-----thalsq-----pmlsiseasdnqjseidfsqjnv 602

QY 381 TKHD-OG-----MGWG-TLDEDPDQGHMR-----FIPL-RHGKYMV 414

DB 381 tkhd-og-----mgwg-tldeppdqghmr-----fipl-rhgkymv 414

QY 603 phygwgqglwtwgtaktqdpapasatitcdpqkanfhtllltwlpagypvapkhrppl 662

DB 603 phygwgqglwtwgtaktqdpapasatitcdpqkanfhtllltwlpagypvapkhrppl 662

OY 415 SSKRPMWFMYMESSASGYIRSMENNPQGHNSIT 450
 Db 663 antlwgmmllateslkn-----saeltspgphfvgit 694

RESULT 3

AA83202
 ID AAG83202 standard; Protein; 964 AA.

XX AAG83202;

XX 05-SEP-2001 (first entry)

XX Protein encoded by Chlamydia trachomatis pmpe gene.

KW Chlamydia; vaccine; infection; fusion protein; antigen;
 KW pelvic inflammatory disease; trachoma; atherosclerosis; heart disease;
 KW acute respiratory tract infection; Capl; CT529; OMCB;
 KW polymorphic membrane protein; pmp; thiol specific antioxidant; TSA.

XX Chlamydia trachomatis.

XX WO200140474-A2.

XX 07-JUN-2001.

XX 04-DEC-2000; 2000WO-US32919.

XX 03-DEC-1999; 99US-0454684.

XX 19-APR-2000; 2000US-0556877.

XX 20-JUN-2000; 2000US-0598419.

XX (CORI-) CORIXA CORP.

XX Probst P, Bhatia A, Skeiky YAW, Fling SP, Scholler J;

XX WPI; 2001-374831/39.

XX Chlamydia polypeptides and fusion proteins useful for preventing pelvic
 PT inflammatory disease, trachoma, acute respiratory tract infections,
 XX atherosclerosis and heart disease -

XX Claim 2; Page 191-193; 295pp; English.

CC The present sequence is provided in a specification relating to
 CC compounds and methods for the treatment and diagnosis of chlamydial
 CC infection. The compounds provided include polypeptides and fusion
 CC proteins comprising immunogenic portions of Chlamydia antigens
 CC and DNA sequences encoding such polypeptides. They are useful for
 CC vaccinating against chlamydial infection, which causes pelvic
 CC inflammatory disease, trachoma, acute respiratory tract infections,
 CC atherosclerosis and heart disease.

XX Sequence 964 AA;

Query Match 4.8%; Score 112.5; DB 22; Length 964;
 Best Local Similarity 21.3%; Pred. No. 0.66;
 Matches 110; Conservative 61; Mismatches 184; Indels 161; Gaps 28;

OY 21 GKSDV-----NSLLTKVETLAKESGSENALEBLEKKEIQTKPRVQAKRIILGSVC 76
 Db 254 gngnvtvfyngnfknvetassas-----dggalkvtctridvtgngrrlffs-- 301
 OY 77 SALGKINSQDATKIISSGLDIVAGIATTEGPGVMGIGAVASFSSILSLFTGSSAKNSV 136
 Db 302 -----dnltknvgal---yaprvltvdgptyffimianhng 336
 OY 137 AAVIDRALSKHRDEAIORNA-----AGAKR-----DPAESSAFYQV 172
 Db 337 galyldgtstnkslisa-drlaiflnenlvtnvtangstsaamprrnatlvasssgelll 395
 OY 173 MKQGS-NLTFUSD-LSITAAVVPYKFSNFIGLESRTISQGAATTSLSDAKRAVD----- 224

Db 396 gaggsgnllfydpdtevsnaevsv-sfkheadqgsvvfgatvnsaafhnrlnqtkpap 454
 OY 225 -----FILL--YQULVMBETL--LVLD--AII--YRKGAHEHVASAVENANRYNKE 268
 Db 455 ltlsgflctedhaqltvtncftgtgvslyngavlsctyngcdg----sasnaslkh 510
 OY 269 LAADTDLFLHKLPEQALLCAVYHPISASETSAILNTY----KYECVPVPRPI-----G 320
 Db 511 lgin---lssllksgeel-----pllwepetnmsnvtadtaetlsldvsksllddyg 561
 OY 321 NRRYKFTNSYWNNTYSICSEAYMGNYMRCGSNVRNPNIRVSKMSDGYTENSDDRRLYI 380
 Db 562 nspyestdl---thalsqg-----pmlsiseasdnqigsenldfsglnv 602
 OY 381 TKHD-QG--WGC-FLDEDPDGDGHR-----FIPL-RHGKRYV 414
 Db 603 phygwgglwvgyaktdqpepassatitdpqkanrfrtlllwpagypspkhrppl 662
 OY 415 SSKRPMWFMYMESSASGYIRSMENNPQGHNSIT 450
 Db 663 antlwgmmllateslkn-----saeltspgphfvgit 694

RESULT 4

AA813640
 ID AAB13640 standard; Protein; 977 AA.

XX AAB13640;

XX 02-FEB-2001 (first entry)

XX C. trachomatis pmpe gene amino terminus minus signal sequence protein.

XX Chlamydial infection; sexually transmitted disease;
 KW pelvic inflammatory disease; PID; tubal obstruction; infertility;
 KW trachoma; blindness; acute respiratory tract infection;
 KW atherosclerosis; coronary heart disease; antibacterial.

XX Chlamydia trachomatis.

XX WO200034483-A2.

XX 15-JUN-2000.

XX 08-DEC-1999; 99WO-US29012.

XX 08-DEC-1998; 98US-0208277.

XX 08-APR-1999; 99US-0208594.

XX 01-OCT-1999; 99US-0410568.

XX 22-OCT-1999; 99US-0426571.

XX (CORI-) CORIXA CORP.

XX Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;

XX WPI; 2000-431303/37.

PT Isolated polypeptide for diagnosis and treatment of Chlamydia infection
 PT comprises immunogenic portion of Chlamydia antigen, which comprises
 PT amino acid sequence encoded by polynucleotide sequence -
 XX Claim 2; Pages 210-212; 256pp; English.

CC The present invention relates to new nucleic acid sequences and the
 CC proteins encoded by the nucleic acid sequences. The encoded proteins
 CC comprise an immunogenic portion of a Chlamydia antigen. The encoded
 CC proteins are useful for the serodiagnosis and treatment of Chlamydia
 CC infection. Chlamydiae are intracellular bacterial pathogens that are
 CC responsible for a wide variety of human infections. C. trachomatis
 CC infection is one of the most common sexually transmitted diseases and can
 CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction
 CC and infertility. Trachoma due to ocular infection with C. trachomatis is

CC the leading cause of preventable blindness worldwide. C. pneumonia is a
CC major cause of acute respiratory tract infections in humans and is also
CC thought to play a role in the pathogenesis of atherosclerosis and
CC coronary heart disease. The present sequence is a protein isolated in the
CC present invention.

XX Sequence 977 AA:

Query Match 4.8%; Score 112.5; DB 21; Length 977;

Best Local Similarity 21.3%; Pred. No. 0.68; Matches 110; Conservative 61; Mismatches 184; Indels 161; Gaps 28;

```
QY 21 GKRSDV-----NSLLRVERALKEASGNEALALEGKELQTKPDRGQATKILGSYG 76
DB 267 gnrgnivfynrcfknvetasseas-----dggalkvtrldvgngrlffs-- 314
QY 77 SALGRLNSGDATKIISGCLDIYAGIATTFGPGVMGIGAVASVSSILSLFTGSSAKNSV 136
DB 315 -----dnltknysgal---yapvvlvdngpctyflnnlanhky 349
QY 137 AAVIDRALSKHRDEAIQRHA-----AGAKR-----DFAESSAFIOY 172
DB 350 gailyldgtnsksisa-drnailfneivnvnangstscanprrnaaltvasssgeill 408
QY 173 MKQOS-NLTDSO-LSIIAANVPYKFSNFIGOLESRIISOGAATTSLSDAKRAVD----- 224
DB 409 gagssgnllfydpievsnagvsv-sfnkeadqtsvsvfsatvnsadffqrmldqtktpap 467
QY 225 -----FILL--YCOLVVMRETL---LVDL---AIL--YRKGAHEVAVASAVENANRVNKE 268
DB 468 ltlslngflcliedhaqltvnrtftgtgvsalngavlsckngtgd-----sasnasitlkh 523
QY 269 LAADTLDFLHKLIPEQALIGAVYHPISASETSKAILNTY---KYRGVDPVPRPI-----G 320
DB 524 lglh-----lssilksgael-----pllwepclnsmnyladaatfsicvklsllddyg 574
QY 321 NRRYKFTSNYWNYSICSEAYMGNYMFRGCSNVNRPNIHVSKMSDGFYTMESDRKRLYI 380
DB 575 nsypsestdl---thalsgq-----pmlsiseasdnqgsenldfsglnv 615
QY 381 TKHD-OG---WGWG-TLDEDPDGOGHR-----FIPL-RHGKYWV 414
DB 616 phygwgglwtwqaktqdppepassatldpqrkanfrtllltwlpagyvspkhrsp1l 675
QY 415 SSKRWPWFMYMESSASGYIRSMENNPQGGHMSIT 450
DB 676 antlwgmmllateslkn-----saeltspshpftwgit 707
```

RESULT 5
AAG83208
ID AAG83208 standard; Protein; 977 AA.

XX AAG83208;

XX 05-SEP-2001 (first entry)

XX Protein encoded by Chlamydia trachomatis pmpe gene.

XX Chlamydia; vaccine; infection; fusion protein; antigen;
XX pelvic inflammatory disease; trachoma; atherosclerosis; heart disease;
XX acute respiratory tract infection; Cap1; CT529; OMBB;
XX polymorphic membrane protein; pmp; thol specific antioxidant; TSA.

XX Chlamydia trachomatis.

XX WO20010474-A2.

XX 07-JUN-2001.
XX 04-DEC-2000; 2000WO-US32919.

PR 03-DEC-1999; 99US-0454684.
PR 19-APR-2000; 2000US-0556877.
PR 20-JUN-2000; 2000US-0598419.

XX (CORI-) CORIXA CORP.

XX Probst P, Bhatia A, Skelky YAW, Fling SP, Scholler J;

XX WPI: 2001-374831/39.

XX Chlamydia polypeptides and fusion proteins useful for preventing pelvic
XX inflammatory disease, trachoma, acute respiratory tract infections,
XX atherosclerosis and heart disease -

XX Claim 2; Page 216-218; 295pp; English.

XX The present sequence is provided in a specification relating to
XX compounds and methods for the treatment and diagnosis of chlamydial
XX infection. The compounds provided include polypeptides and fusion
XX proteins comprising immunogenic portions of Chlamydia antigens
XX and DNA sequences encoding such polypeptides. They are useful for
XX vaccinating against chlamydial infection, which causes pelvic
XX inflammatory disease, trachoma, acute respiratory tract infections,
XX atherosclerosis and heart disease.

XX Sequence 977 AA:

Query Match 4.8%; Score 112.5; DB 22; Length 977;

Best Local Similarity 21.3%; Pred. No. 0.68; Matches 110; Conservative 61; Mismatches 184; Indels 161; Gaps 28;

```
QY 21 GKRSDV-----NSLLRVERALKEASGNEALALEGKELQTKPDRGQATKILGSYG 76
DB 267 gnrgnivfynrcfknvetasseas-----dggalkvtrldvgngrlffs-- 314
QY 77 SALGRLNSGDATKIISGCLDIYAGIATTFGPGVMGIGAVASVSSILSLFTGSSAKNSV 136
DB 315 -----dnltknysgal---yapvvlvdngpctyflnnlanhky 349
QY 137 AAVIDRALSKHRDEAIQRHA-----AGAKR-----DFAESSAFIOY 172
DB 350 gailyldgtnsksisa-drnailfneivnvnangstscanprrnaaltvasssgeill 408
QY 173 MKQOS-NLTDSO-LSIIAANVPYKFSNFIGOLESRIISOGAATTSLSDAKRAVD----- 224
DB 409 gagssgnllfydpievsnagvsv-sfnkeadqtsvsvfsatvnsadffqrmldqtktpap 467
QY 225 -----FILL--YCOLVVMRETL---LVDL---AIL--YRKGAHEVAVASAVENANRVNKE 268
DB 468 ltlslngflcliedhaqltvnrtftgtgvsalngavlsckngtgd-----sasnasitlkh 523
QY 269 LAADTLDFLHKLIPEQALIGAVYHPISASETSKAILNTY---KYRGVDPVPRPI-----G 320
DB 524 lglh-----lssilksgael-----pllwepclnsmnyladaatfsicvklsllddyg 574
QY 321 NRRYKFTSNYWNYSICSEAYMGNYMFRGCSNVNRPNIHVSKMSDGFYTMESDRKRLYI 380
DB 575 nsypsestdl---thalsgq-----pmlsiseasdnqgsenldfsglnv 615
QY 381 TKHD-OG---WGWG-TLDEDPDGOGHR-----FIPL-RHGKYWV 414
DB 616 phygwgglwtwqaktqdppepassatldpqrkanfrtllltwlpagyvspkhrsp1l 675
QY 415 SSKRWPWFMYMESSASGYIRSMENNPQGGHMSIT 450
DB 676 antlwgmmllateslkn-----saeltspshpftwgit 707
```

RESULT 6
AAG49832
ID AAG49832 standard; Protein; 1066 AA.

AC AAG49832;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 63081.
XX
KW Protein identification: signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP103405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148177.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149366.
PR 17-AUG-1999; 99US-0149366.
PR 18-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.

PR 26-AUG-1999;	99US-0150884.
PR 27-AUG-1999;	99US-0151065.
PR 27-AUG-1999;	99US-0151066.
PR 27-AUG-1999;	99US-0151080.
PR 30-AUG-1999;	99US-0151303.
PR 31-AUG-1999;	99US-0151438.
PR 01-SEP-1999;	99US-0151930.
PR 07-SEP-1999;	99US-0153363.
PR 10-SEP-1999;	99US-0153070.
PR 13-SEP-1999;	99US-0153758.
PR 15-SEP-1999;	99US-0154018.
PR 16-SEP-1999;	99US-0154039.
PR 20-SEP-1999;	99US-0154779.
PR 22-SEP-1999;	99US-0155139.
PR 23-SEP-1999;	99US-0155486.
PR 24-SEP-1999;	99US-0155695.
PR 28-SEP-1999;	99US-0156458.
PR 29-SEP-1999;	99US-0155596.
PR 04-OCT-1999;	99US-0157117.
PR 05-OCT-1999;	99US-0157753.
PR 06-OCT-1999;	99US-0157865.
PR 07-OCT-1999;	99US-0158029.
PR 08-OCT-1999;	99US-0158232.
PR 12-OCT-1999;	99US-0158369.
PR 13-OCT-1999;	99US-0159293.
PR 13-OCT-1999;	99US-0159294.
PR 13-OCT-1999;	99US-0159295.
PR 14-OCT-1999;	99US-0159329.
PR 14-OCT-1999;	99US-0159330.
PR 14-OCT-1999;	99US-0159331.
PR 14-OCT-1999;	99US-0159637.
PR 14-OCT-1999;	99US-0159638.
PR 18-OCT-1999;	99US-0159584.
PR 21-OCT-1999;	99US-0160741.
PR 21-OCT-1999;	99US-0160767.
PR 21-OCT-1999;	99US-0160768.
PR 21-OCT-1999;	99US-0160770.
PR 21-OCT-1999;	99US-0160814.
PR 21-OCT-1999;	99US-0160815.
PR 22-OCT-1999;	99US-0160980.
PR 22-OCT-1999;	99US-0160981.
PR 22-OCT-1999;	99US-0160989.
PR 25-OCT-1999;	99US-0161404.
PR 25-OCT-1999;	99US-0161405.
PR 25-OCT-1999;	99US-0161406.
PR 26-OCT-1999;	99US-0161359.
PR 26-OCT-1999;	99US-0161360.
PR 26-OCT-1999;	99US-0161361.
PR 28-OCT-1999;	99US-0161920.
PR 28-OCT-1999;	99US-0161992.
PR 28-OCT-1999;	99US-0161993.
PR 29-OCT-1999;	99US-0162142.

	Query Match	4.7%	Score 110;	DB 21,	length 1066;	
	Best Local Similarity	22.8%;	Pred. No. 1.3;			
	Matches	93;	Conservative	51;	Mismatches	118; Indels 146; Gaps 20;
OY	9 LFVYLAIRSAHKGRSD-----VNSLTKKETALKEASGSNEALELEGLK	55				
	: : : : :		: : : : :			
Db	190 llllaavaslslgltlgeglkeglldgsiafaflvllivrlavsdyrqslq-fqnIndex	247				
OY	56 GEIQTKPPRVCQATKI-----LSVSALGKLMSGATKI-----	90				
	: : : : :		: : : : :			
Db	248 nrlqlvemrgrrtvklslsydvvgdvlprlrldgdpadgvlislghalddssmtgeskl	307				
OY	91 -----TSGC-----DLIVGIATFTFG-----GPVMGTGAVAS	118				
	: : : : :		: : : : :			
Db	308 vhhkqdkpfmsgcgvadvvgumlvtygintewglmasedtgeeeplygringlit	367				
OY	119 FVSSILSLFTGGSSAKNSVAAYVDRLSKRHDEAIOHHAAGAKRDFEESAFIQWKOQN	178				
	: : : : :		: : : : :			
Db	368 fl-givgl-----svallvllval-----lvryftgtctgdengatqfiktstisd	411				

Oy	179	LTDG-----DLSTIANVPE-----VYKFSNFTGO--LESRIIS----OGAATP 214
Dd	412	Ivdcvkqftflavtlvwvaapeglplavtlltlaysmikmmdakalvirrlsacetsgatl 471
Oy	215	SLSDAKRAAVDEILLYCOLVMRETLVLVDLALVYRKGAENHVASAVENANRVNELADTL 274
Dd	472	Iscdktyc---ltnqlwtv-et-----yagsgkmdva---dnpsglhpklva--- 512
Oy	275	DELKLRIPE---QALIGAVHP-----TSASTSKAILNTYKRFV 312
Dd	513	-----lisegvagnttgnilmpkdgeveisgsplekallswaykkigm 555
 RESULT 7 AAC49831 ID AAC49831 standard; Protein; 1073 AA. AC AC XX AAC49831; XX DT 18-OCT-2000 (first entry) XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 63080. XX KW Protein identification; signal transduction pathway; metabolic pathway; KW hydrolisation assay; genetic mapping; gene expression control; promoter XX termination sequence. XX OS Arabidopsis thaliana. XX EN EP1033405-AZ. XX PD 06-SEP-2000. XX PF 25-FEB-2000; 2000EP-0301439. XX PR 25-FEB-1999; 99US-0121825. PR 05-MAR-1999; 99US-0123180. PR 09-MAR-1999; 99US-0123548. PR 23-MAR-1999; 99US-0125788. PR 25-MAR-1999; 99US-0126264. PR 29-MAR-1999; 99US-0126785. PR 01-APR-1999; 99US-0127462. PR 06-APR-1999; 99US-0128234. PR 08-APR-1999; 99US-0128714. PR 16-APR-1999; 99US-0129845. PR 19-APR-1999; 99US-0130077. PR 21-APR-1999; 99US-0130449. PR 23-APR-1999; 99US-0130510. PR 23-APR-1999; 99US-0130891. PR 28-APR-1999; 99US-0131449. PR 30-APR-1999; 99US-0132048. PR 30-APR-1999; 99US-0132407. PR 04-MAY-1999; 99US-0132484. PR 05-MAY-1999; 99US-0132485. PR 06-MAY-1999; 99US-0132486. PR 06-MAY-1999; 99US-0132487. PR 07-MAY-1999; 99US-0132863. PR 11-MAY-1999; 99US-0134256. PR 14-MAY-1999; 99US-0134218. PR 14-MAY-1999; 99US-0134219. PR 14-MAY-1999; 99US-0134221. PR 14-MAY-1999; 99US-0134370. PR 18-MAY-1999; 99US-0134768. PR 19-MAY-1999; 99US-0134941. PR 20-MAY-1999; 99US-0135124. PR 21-MAY-1999; 99US-0135353. PR 24-MAY-1999; 99US-0135629. PR 25-MAY-1999; 99US-0136021. PR 27-MAY-1999; 99US-0136392. PR 28-MAY-1999; 99US-0136782. PR 01-JUN-1999; 99US-0137222. PR 03-JUN-1999; 99US-0137528. PR 04-JUN-1999; 99US-0137502.		

PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142802.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150864.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 4.78; Score 110; DB 21; Length 1073;
Best Local Similarity 22.88; Pred. No. 1.3;
Matches 93; Conservative 51; Mismatches 118; Indels 146; Gaps 20;

[illegible][illegible]

```

PR      22-OCT-1999;          99US-0160980.
PR      22-OCT-1999;          99US-0160981.
PR      22-OCT-1999;          99US-0160981.
PR      22-OCT-1999;          99US-0160989.
PR      25-OCT-1999;          99US-0161404.
PR      25-OCT-1999;          99US-0161405.
PR      25-OCT-1999;          99US-0161407.
PR      25-OCT-1999;          99US-0161406.
PR      26-OCT-1999;          99US-0161359.
PR      26-OCT-1999;          99US-0161360.
PR      26-OCT-1999;          99US-0161361.
PR      28-OCT-1999;          99US-0161920.
PR      28-OCT-1999;          99US-0161921.
PR      28-OCT-1999;          99US-0161993.
PR      29-OCT-1999;          99US-0162142.

Query Match
Best Local Similarity    4.7%; Score 110; DB 21; Length 1086;
Matches   93; Conservative   51; Mismatches 118; Indels 146; Gap

QY      9 LPIVLAITSAKHGKRS-----VNSLTKRVETALEAGSNALEAEGLK 5
DB      210 llllaavetslalgiktgeglkgldgsiafavlllvltavaaydrgslq-fgnIndex 2
QY      56 GEIQTKPPRVOGATKI-----IGSVGSALGKLNSGDATKI----- 90
DB      268 rnlqdlewmrgrlrvtvkkslydvvgdviprlrdgdvpadgvlisghsladesmtgeski 32
QY      91 -----ISGC-----LDIVAGIATPFG-----GPVGMGICGVAS 11
DB      328 vnhdqgkspflmgcgkvadvgvnmllvlgvyintcwgjllmasisedtgetelplrvrlnglat 387
QY      119 FVSJILSLTGGSSANKSNVAVIDRALSKIRDEAIORHAAGAKRDFAESSAFIQVMKOOSN 178
DB      388 fl-glvgl-----swalvvivaal-----lvryftgtqtqdgatgfikytsisd 431
QY      179 LTDS-----DLSTIAANVP-----VKFSPNFIO--LESRTS---OGAATT 214
DB      432 lvddcvkhfllaavlivvaapegiplaetllclaysmkmmackalvirrisaceltsgsat 491
QY      215 SLSDAKRAVDFIILLCOLVYMRETLAVDAILFKRGNAEHVASAVENANRYNKELADTL 274
DB      492 tscdkrtgt-----tlngmvtv-et-----yagsskmgya---dnpsglhpklva--- 532
QY      275 DELHKLIPE---OALIGAYVR-----IASETSKAILENTYRKYGCV 312
DB      533 ----lllseqvagnntgnlfhpkdggveaisgspekallsiwayxlgw 575

RESULT       9
AA127230
ID AAY27230 standard; Protein; 1978 AA.
AC AAY27230;
XX XX
DT 24-SEP-1999 (first entry)
DE Amino acid sequence of N. meningitidis protein ORF114-1.
KW Neisseria meningitidis protein; pharmaceutical; vaccine; diagnosis;
KM bacterial infection; treatment.
OS Neisseria meningitidis.
PN W09936544-A2.
PD 22-JUL-1999.
PF 14-JAN-1999;     99MO-IB00103.
PR 09-OCT-1998;     98GB-0022143.
PR 14-JAN-1998;     98CB-0002143.
```


Db 713 ltlmalegfhvnglqghaslllaal-----kdgalvtvdagammkeesllqmaa 764
QY 158 G-AKRD-----PAESSAFIOVMKQOONLTDSDLSIIAANVPYKFS---NFIGOLESR1 207
Db 765 nqekdlkltswtqidaalqwlqmsalavspidl--agmalkygidhnyaaaw----- 817
QY 208 SGAATTSLSL-----AKRAVD--FLLYCOLVVRRETLVDLAILYRKGAHEHVASAVEN 261
Db 818 -gaaaaalmadhangqkklidelfskalcnyl---navdsaaagvrdnnglytllidn 873
QY 262 -----ANRVKELADDTLDFLHKLIPDQALIGAVYHPISASETSKAILNTKRYGVPD 314
Db 874 qvsadvltstriaaiaaglylvnralnrdegqla-----sdvstrqfftdwery----- 922
QY 315 VPRPIGNRRKFTNSYWNYSICSEA--YMGNNMFRGCSNVRNPINRV--SKMSDGFYTM 370
Db 923 -----nkry-----stwagvselvyypeny-----dptgrlqgkmmadallqs 961
QY 371 EN-----SDRRKLYITKHDGKMGCTL-----DEDPDQGHMRFPL----- 407
Db 962 lngsqlnadltvedafktyltsfegvanlkvisayhndvndqgltyfigidgaapgytw 1021
QY 408 -----RHGKYVSSKRRPNNFMWMESSASGYIRSMEN 439
Db 1022 rsvdshkscengkf--aaanagew-----nkltcavnpwkn 1054

RESULT 11

AAW56558
ID AAW56558 standard; Protein; 1844 AA.

AC AAW56558;

DT 07-AUG-1998 (first entry)

XX Toxin TcdAII, encoded by the tcdA gene from genomic region tcd.

XX Photorhabdus luminescens W-14; nematode; symbiotic;

XX Heterorhabdus; tca; tcb; tcc; tcd; insecticidal activity; toxin;

XX Lepidoptera; Coleoptera; Hymenoptera; Diptera; Dictyoptera; Acarina;

XX Homoptera; Southern; Western corn rootworm; Colorado potato beetle;

XX mealworm; boll weevil; turf grub; beetle armyworm; black cutworm;

XX cabbage looper; codling moth; corn earworm; European corn borer;

XX tobacco hornworm; budworm.

XX Photorhabdus luminescens.

XX W09080932-A1.

XX 05-MAR-1998.

XX 05-MAY-1997; 97MO-US07657.

XX 06-NOV-1996; 96WO-US18003.

XX 28-AUG-1996; 96US-0705484.

XX 06-NOV-1996; 96US-0743699.

XX (DOMC) DOWELANCO.

XX (WISC) WISCONSIN ALUMNI RES FOUND.

XX Blackburn MB, Bowen DJ, Cliche JA, Ensign JC, Fatig R;
XX French-Constant RH, Guo L, Hey JD, Merlo DJ, Orr GU;
XX Petrelli J, Roberts JL, Rochelleau TA, Schoonover S;
XX Strickland JA, Sukhplinda K;
XX MPI, 1998-179427/16.
XX N-PSDB; AAV29986.

CC The present sequence represents a protein named TcdAII of the bacterium
CC Photorhabdus luminescens (W-14). This is a symbiotic bacterium of the
CC nematodes of the Heterorhabdus genus. The bacterium has at least 4
CC distinct genomic regions, tca, tcb, tcc, and tcd. Peptide products are
CC produced from these regions that are associated with insecticidal
CC activity. Peptides AAW56560 and AAW56562-65 are fragments of the present
CC protein. The native toxins are secreted proteins. The proteins are
CC toxic to insects upon exposure and especially when ingested. The
CC nucleic acid sequence can be used to produce transgenic plants,
CC baculoviruses or microbial hosts for toxin production. They can be used
CC to control insects pests from the Lepidoptera, Coleoptera, Hymenoptera,
CC Diptera, Dictyoptera, Acarina or Homoptera orders, especially the
CC Southern or Western corn rootworm, Colorado potato beetle, mealworm,
CC boll weevil, turf grub, beetle armyworm, black cutworm, cabbage looper,
CC codling moth, corn earworm, European corn borer or tobacco hornworm
CC or budworm.

CC Sequence 1844 AA;

Query Match 4.5%; Score 105.5; DB 19; Length 1844;
Best Local Similarity 19.2%; Pred No. 7.2; Mismatches 184; Gaps 27;
Matches 100; Conservative 85; Indels 151;

QY 11 IYLAITSKHKGRS---DVNSLTKRVETALKKASGSNEALEAL-----EG 53
Db 595 ltatstslhgkesllggedlkrampcftsalht-sgvevdl1llwidqpaqltvdg 653
QY 54 LKGEIOTKPRDR-----GQATKILGVSALGKLSGDKATKISCLDIVAGIATFGGP 108
Db 654 fweevqtptslkvltfagvlqajsllyrrlg-lsetelslvtqsl1lvagksllhdgl 712
QY 109 VGM-----GIGAVASFVSSILFLFGSSKKNVAAYIDALSKHREAIQRNA 157
Db 713 ltlmalegfhvnglqghaslllaal-----kdgalvtvdagammkeesllqmaa 764
QY 158 G-AKRD-----PAESSAFIOVMKQOONLTDSDLSIIAANVPYKFS---NFIGOLESR1 207
Db 765 nqekdlkltswtqidaalqwlqmsalavspidl--agmalkygidhnyaaaw----- 817
QY 208 SGAATTSLSL-----AKRAVD--FLLYCOLVVRRETLVDLAILYRKGAHEHVASAVEN 261
Db 818 -gaaaaalmadhangqkklidelfskalcnyl---navdsaaagvrdnnglytllidn 873
QY 262 -----ANRVKELADDTLDFLHKLIPDQALIGAVYHPISASETSKAILNTKRYGVPD 314
Db 874 qvsadvltstriaaiaaglylvnralnrdegqla-----sdvstrqfftdwery----- 922
QY 315 VPRPIGNRRKFTNSYWNYSICSEA--YMGNNMFRGCSNVRNPINRV--SKMSDGFYTM 370
Db 923 -----nkry-----stwagvselvyypeny-----dptgrlqgkmmadallqs 961
QY 371 EN-----SDRRKLYITKHDGKMGCTL-----DEDPDQGHMRFPL----- 407
Db 962 lngsqlnadltvedafktyltsfegvanlkvisayhndvndqgltyfigidgaapgytw 1021
QY 408 -----RHGKYVSSKRRPNNFMWMESSASGYIRSMEN 439
Db 1022 rsvdshkscengkf--aaanagew-----nkltcavnpwkn 1054

RESULT 12
ID AAW17871
AAW17871 standard; Protein; 2504 AA.
AC AAW17871;
DT 29-JAN-1998 (first entry)
XX Photorhabdus luminescens insect toxin protein TcdA.
XX Insecticide; insect; pest control; biological control;
XX Photorhabdus luminescens; TcdA; Southern corn rootworm;

KM Colorado potato beetle; Western corn rootworm; meal worm;
 KM boll weevil; turf grub; Coleoptera; beet armyworm; black cutworm;
 KM cabbage looper; codling moth; corn earworm; European corn borer;
 KM tobacco hornworm; tobacco budworm; Lepidoptera; Hymenoptera;
 KM Diptera, Dictyoptera; Acarina; Homoptera.
 OS Photorhabdus luminescens strain W-14 (ATCC 55397).
 XX
 PM WO9117432-A1.
 XX
 PD 15-MAY-1997.
 XX
 PF 06-NOV-1996; 96WO-US18003.
 XX
 PR 28-AUG-1996; 96US-0705484.
 PR 06-NOV-1995; 95US-0007255.
 PR 28-FEB-1996; 96US-0608423.
 XX
 PA (WISC) WISCONSIN ALUMNI RES FOUND.
 XX
 PI Blackburn MB, Bowen DJ, Cliche TA, Ensign JC, Fatig R;
 PI Ffrench-Constant RH, Guo L, Hey TD, Merlo DJ, Orr GL;
 PI Petrell J, Roberts JL, Rocheleau TA, Schoonover S;
 PI Strickland JA.
 XX
 DR WPI. 1997-281022/25.
 DR N-PSDB; AAT68836.
 XX
 PT Photorhabdus sp. insecticidal protein toxins and DNA encoding them -
 PT can be genetically engineered into insect larvae food and plants for
 PT insect control
 XX
 PS Claim 8; Page 119-129; 276pp; English.
 XX
 CC This polypeptide comprises a specifically claimed 280.6 kDa
 CC insecticidal toxin, Tcda, of Photorhabdus luminescens that is
 CC proteolytically cleaved into 207.6 kDa (see AAM18302) and 62.9 kDa
 CC polypeptides (see AAM18303). The Tcda polypeptide can be expressed
 CC in host cells using a gene (see AAT68836) isolated from a genomic
 CC library. Claimed toxin proteins of P. luminescens (see AAM17871,
 CC AAM17884-89, AAM17899-900, AAM18301-06) can be applied to, or
 CC genetically engineered into, insect larvae food and plants for insect
 CC control. The Photorhabdus toxins are particularly effective against
 CC Southern corn rootworm, Colorado potato beetle, Western corn rootworm,
 CC meal worm, boll weevil and turf grub (Coleoptera), beet armyworm, black
 CC cutworm, cabbage looper, codling moth, corn earworm, European corn
 CC borer, tobacco hornworm and tobacco budworm (Lepidoptera), and
 CC are also active against insects of the orders Hymenoptera, Diptera
 CC Dictyoptera, Acarina and Homoptera. (All claimed).
 CC
 XX
 SQ Sequence 2504 AA;
 Query Match 4.5%; Score 105.5; DB 18; Length 2504;
 Best Local Similarity 19.2%; Pred. No. 11;
 Matches 100; Conservative 85; Mismatches 184; Indels 151; Gaps 27;
 OY 11 IYLAITSAKHGRS---DVNSLTKVETALKKASGSENALEAI-----EG 53
 DB 682 ltaiststlhgkeslglgelkxramapcfsalhtlsgvaydlilwddqpaqitvdg 740
 OY 54 LKGEIQTAKPDY-----GATKIIGSVGALGKLSGDAKTKITSCGLDIVAGIATTFGCP 108
 DB 741 fweevgtiptslkvlftfaqlsllyrrtlylsetelslvtssllvaggksldhql 799
 OY 109 VGM-----GIGAVASFVSSILSLFTSSAKNSVAAYIDRALSKHREAIORHNA 157
 DB 800 lllmalegfhtwnglqgqaslllaa-----kdgalvtvtvgaqamkkesllqmaa 851
 OY 158 G-AKRD-----FAESSAFIQVMKQOQNSLTSDSLTIAANVPYKFS---NFIGQLESRI 207
 DB 852 ngvexdkltktsawtgdailqylqmsalavspldl--agmalxyglahnyaaaw----- 904

OY 208 SGAATSTLSD---AKRAVD--FILLYCOLVYMRRETLVDLAILYRKGNAEHVASAVEN 261
 DB 905 -qaaaalmadanqaqkldetlskalcnyl---navdsaaqvtrnglyl1idn 960
 OY 262 -----ANRYNKELEAATDLELHKLIGAUVYHPIASSETSKA1LWTKYFGYVD 314
 DB 961 qvsadvtsrltaeaaqaglylvrnalndegla-----sdvstrgfctwery----- 1009
 OY 315 VPRPIGNRRYFTNSYVWNTYSICSEA--YMGNYMERGCSNVNRPNRY--SKMSDGEYTM 370
 DB 1010 -----nkry-----stwagvelvyypenyv-----dprqrlgqtkmmdallgs 1048
 OY 371 EN-----SDRRKLYTKHNDGWMGTL----DEDPGOGHNRFTPL----- 407
 DB 1049 lngsqlnadvedafkkytlfefegvaanlkvisayhndvndqgllyflgldgaapgyw 1108
 OY 408 -----RHGKYVWSKRWPMFWMFYMESSASGYIRSMEN 439
 DB 1109 rsvdskcengkf--aanawgew-----nkltcavnpkkn 1141
 RESULT 13
 ID AAM56557 standard; Protein: 2504 AA.
 XX
 AC AAM56557;
 XX
 DT 07-AUG-1998 (first entry)
 XX
 DE Toxin Tcda, encoded by the tcda gene from genomic region tcb.
 XX
 KW Photorhabdus luminescens W-14; nematode; symbiotic;
 KW Heterorhabditis; tca; tcb; tcc; tcd; insecticidal activity; toxin;
 KW Lepidoptera; Coleoptera; Hymenoptera; Diptera; Dictyoptera; Acarina;
 KW Homoptera; Southern; Western corn rootworm; Colorado potato beetle;
 KW mealworm; boll weevil; turf grub; beetle armyworm; black cutworm;
 KW cabbage looper; codling moth; corn earworm; European corn borer;
 KW Tobacco hornworm; budworm.
 XX
 OS Photorhabdus luminescens.
 XX
 PN WO9808932-A1.
 XX
 PD 05-MAR-1998.
 XX
 PF 05-MAY-1997; 97WO-US07657.
 XX
 PR 06-NOV-1996; 96WO-US18003.
 PR 28-AUG-1996; 96US-0705484.
 PR 06-NOV-1996; 96US-0743699.
 XX
 PA (DOWC) DOWELANCO.
 PA (WISC) WISCONSIN ALUMNI RES FOUND.
 XX
 PI Blackburn MB, Bowen DJ, Cliche TA, Ensign JC, Fatig R;
 PI Ffrench-Constant RH, Guo L, Hey TD, Merlo DJ, Orr GL;
 PI Petrell J, Roberts JL, Rocheleau TA, Schoonover S;
 PI Strickland JA, Sukhapinda K;
 XX
 DR WPI. 1998-179427/16.
 DR N-PSDB; AAV29985.
 XX
 PT Isolated toxins from Photorhabdus luminescens strains - useful for
 PT control of insect pests
 XX
 PS Claim 34; Pages 163-169; 321pp; English.
 CC The present sequence represents a protein named Tcda of the bacterium
 CC Photorhabdus luminescens (W-14). This is a symbiotic bacterium of the
 CC nematodes of the Heterorhabditis genus. The bacterium has at least 4
 CC distinct genomic regions, tca, tcb, tcc, and tcd. Peptide products are
 CC produced from these regions that are associated with insecticidal
 CC activity. The native toxins are secreted proteins. The proteins are

CC toxic to insects upon exposure and especially when ingested. The
 CC nucleic acid sequence can be used to produce transgenic plants.
 CC baculoviruses or microbial hosts for toxin production. They can be used
 CC to control insects pests from the lepidoptera, coleoptera, Hymenoptera,
 CC diptera, Dictyoptera, Acarina or Homoptera orders, especially the
 CC Southern or Western corn rootworm, Colorado potato beetle, mealworm,
 CC boll weevil, turf grub, beetle armyworm, black cutworm, cabbage looper,
 CC codling moth, corn earworm, European corn borer or tobacco hornworm
 CC or budworm.
 CC
 CC
 CC

Sequence 2504 AA:

Query Match 4.5%; Score 105.5; DB 19; Length 2504;
 Best Local Similarity 19.2%; Pred. No. 11;
 Matches 100; Conservative 85; Mismatches 184; Indels 151; Gaps 27;

QY 11 IYLAITSANHGKRS---DVNSLLTKVETALKEASGSENAALRAL-----EG 53
 DB 682 ltaclstlhgkeslgedlkrampcfsalhltsgevaqllllwidqigpaqltvdg 740
 QY 54 LKEIOTKRDV-----GQATKILSGVSGALKLNAGDATKTIISGCLDIYAGIATTFGCP 108
 DB 741 fweevqtptslkvlftaqlaglslyrrig-lsetelslvtgssllvaqkslldhgi 799
 QY 109 VGM-----GIGAVASFEVSSILSFTGSSAKNSVAAYIDRLSKHRDEAIORHAA 157
 DB 800 ltlmalegfhtwvnglqghaslllaal-----kdgaltytdvagammkesllqmaa 851
 QY 158 G-AKRD-----FAESSAFIQVKKQOQSNLTDSDSIITANVPYKFS---NFIGLESRI 207
 DB 852 ngvexdltkltswtqidalqlwlgmsalavspdl-agmalkygidhnyaa-----904
 QY 208 SGGATTSLSLSD---AKRAVD--FILLYQOLVMMRETLVLDALILRKGAHEHVASAVEN 261
 DB 905 -qaaaaalnadhangqakldetfskalcnyy1--navvdsaaayrdnglytllidh 960
 QY 262 -----ANRVKELADTLDFLHKLIPQALIGAVYHPIASSETS KALINTKRYGVPD 314
 DB 961 qvsadvlstrlaealagqllyvnrainrdqgla-----sdvstrqfttdewery-----1009
 QY 315 VRRPIGNRRYKFTNSYWNYSICSEA--YMGNYMERGCSNVRNPNTIRV--SKMSDGFYTM 370
 DB 1010 -----nkry-----stwagvselyvypeny-----dpqtrigqtkmmdallqs 1048
 QY 371 EN-----SDRRKLYITKHDQGWGCTL-----DEDPGDOGHMRFITL-----407
 DB 1049 lngsqlnadvedafkyltsfeyvanlkvlsayhndhvnvdgqlyfigldgaapqlyuw 1108
 QY 408 -----RHGRYVSSSKRPMFMFMYSASGYSIRSMEN 439
 DB 1109 rsvdshkengkf--aanaagew-----nkltcavnpwkn 1141

RESULT 14

AAB72610
 ID AAB72610 standard; Protein: 2504 AA.

AC AAB72610;

XX 04-MAY-2001 (first entry)

DE Photorhabdus tcbd toxin.

XX Tcbd; Insect toxin; plant; insect resistance.

OS Photorhabdus sp.

XX WO200111029-A1.

XX 15-FEB-2001.

PF 11-AUG-2000; 2000WO-US22237.

XX 11-AUG-1999; 99US-0148356.

XX (DOWC) DOW AGROSCIENCES LLC.

XX Petell JK, Melio DJ, Herman RA, Roberts JL, Guo L, Schafer BW;

XX Sukhaphinda K, Melio AO;

XX WPI; 2001-191536/19.

XX N-PSDB; AAF58779.

PT Novel polynucleotide sequence encoding insect toxins, useful for
 PT producing transgenic plants having resistance to insects, especially
 PT corn rootworm
 PS Disclosure; Page 62-72; 106pp; English.

The present invention provides the protein and coding sequences of
 modified versions of the Photorhabdus Tcbd and tcbd toxins. These are
 suitable for expression in plants. The toxins are effective against
 insects upon ingestion, and the sequences provided can be used to produce
 transgenic plants with insect resistance. The present sequence is the
 Photorhabdus tcbd protein.

Sequence 2504 AA:

Query Match 4.5%; Score 105.5; DB 22; Length 2504;
 Best Local Similarity 19.2%; Pred. No. 11;
 Matches 100; Conservative 85; Mismatches 184; Indels 151; Gaps 27;

QY 11 IYLAITSANHGKRS---DVNSLLTKVETALKEASGSENAALRAL-----EG 53
 DB 682 ltaclstlhgkeslgedlkrampcfsalhltsgevaqllllwidqigpaqltvdg 740
 QY 54 LKEIOTKRDV-----GQATKILSGVSGALKLNAGDATKTIISGCLDIYAGIATTFGCP 108
 DB 741 fweevqtptslkvlftaqlaglslyrrig-lsetelslvtgssllvaqkslldhgi 799
 QY 109 VGM-----GIGAVASFEVSSILSFTGSSAKNSVAAYIDRLSKHRDEAIORHAA 157
 DB 800 ltlmalegfhtwvnglqghaslllaal-----kdgaltytdvagammkesllqmaa 851
 QY 158 G-AKRD-----FAESSAFIQVKKQOQSNLTDSDSIITANVPYKFS---NFIGLESRI 207
 DB 852 ngvexdltkltswtqidalqlwlgmsalavspdl-agmalkygidhnyaa-----904
 QY 208 SGGATTSLSLSD---AKRAVD--FILLYQOLVMMRETLVLDALILRKGAHEHVASAVEN 261
 DB 905 -qaaaaalnadhangqakldetfskalcnyy1--navvdsaaayrdnglytllidh 960
 QY 262 -----ANRVKELADTLDFLHKLIPQALIGAVYHPIASSETS KALINTKRYGVPD 314
 DB 961 qvsadvlstrlaealagqllyvnrainrdqgla-----sdvstrqfttdewery-----1009
 QY 315 VRRPIGNRRYKFTNSYWNYSICSEA--YMGNYMERGCSNVRNPNTIRV--SKMSDGFYTM 370
 DB 1010 -----nkry-----stwagvselyvypeny-----dpqtrigqtkmmdallqs 1048
 QY 371 EN-----SDRRKLYITKHDQGWGCTL-----DEDPGDOGHMRFITL-----407
 DB 1049 lngsqlnadvedafkyltsfeyvanlkvlsayhndhvnvdgqlyfigldgaapqlyuw 1108
 QY 408 -----RHGRYVSSSKRPMFMFMYSASGYSIRSMEN 439
 DB 1109 rsvdshkengkf--aanaagew-----nkltcavnpwkn 1141

RESULT 15

AAB72612
 ID AAB72612 standard; Protein: 2505 AA.

AC AAB72612;

XX	04-MAY-2001 (first entry)
DT	
XX	Modified Photornhabdus tcba toxin.
DE	
KW	Tcda; Tcda; insect toxin; plant; insect resistance.
XX	
OS	Photornhabdus sp.
XX	
OS	Synthetic.
FN	WO200111029-A1.
XX	
PD	15-FEB-2001.
XX	
PF	11-AUG-2000; 2000WO-US22237.
XX	
PR	11-AUG-1999; 99US-0148356.
XX	
PA	(DOWC) DOW AGROSCIENCES LLC.
XX	
PI	Petrell JK, Merlo DJ, Herman RA, Roberts JL, Guo L, Schafer BW,
PI	Sukhapinda K, Merlo AO;
XX	
DR	WPI; 2001-191536/19.
XX	
DR	N-PSDB; AAF58781.
XX	
PT	Novel polynucleotide sequence encoding insect toxins, useful for
PT	producing transgenic plants having resistance to insects, especially
PT	corn rootworm
XX	
PS	Claim 1; Page 83-93; 106pp; English.
XX	
CC	The present invention provides the protein and coding sequences of
CC	modified versions of the Photornhabdus Tcda and Tcda toxins. These are
CC	suitable for expression in plants. The toxins are effective against
CC	insects upon ingestion, and the sequences provided can be used to produce
CC	transgenic plants with insect resistance. The present sequence is the
CC	modified Photornhabdus Tcda protein.
XX	
XX	Sequence 2505 AA;
XX	